

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:53:38 ; Search time 6264.22 Seconds
(without alignments)
11470.669 Million cell updates/sec

Title: US-09-610-313-30
Perfect score: 2469
Sequence: 1 gtcgacgccaccatggccga.....gggctagcaccggtaattc 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1:	gb_ba.*	GenEmbl.*
2:	gb_htg.*	
3:	gb_in.*	
4:	gb_om.*	
5:	gb_ov.*	
6:	gb_pat.*	
7:	gb_ph.*	
8:	gb_pl.*	
9:	gb_pr.*	
10:	gb_ro.*	
11:	gb_sts.*	
12:	gb_sy.*	
13:	gb_un.*	
14:	gb_vi.*	
15:	em_ba.*	
16:	em_fun.*	
17:	em_hum.*	
18:	em_in.*	
19:	em_mu.*	
20:	em_om.*	
21:	em_or.*	
22:	em_ov.*	
23:	em_pat.*	
24:	em_ph.*	
25:	em_pi.*	
26:	em_ro.*	
27:	em_sts.*	
28:	em_un.*	
29:	em_vi.*	
30:	em_htg_hum.*	
31:	em_htg_inv.*	
32:	em_htg_other.*	
33:	em_htg_mus.*	
34:	em_htg_pln.*	
35:	em_htg_rod.*	
36:	em_htg_mam.*	
37:	em_htg_vrt.*	
38:	em_sy.*	
39:	em_htgo_hum.*	
40:	em_htgo_mus.*	
41:	em_htgo_other.*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	2469	100.0	2469	6	AX455914	Sequence
2	2442.2	98.9	2463	6	AX455915	Sequence
3	2415.4	97.8	2457	6	AX455916	Sequence
4	2051.2	83.1	9166	6	AX427930	Sequence
5	2042.8	82.7	9788	6	AX427936	Sequence
6	2028.8	82.2	9169	6	AX427931	Sequence
7	2027	82.1	9194	6	AX427925	Sequence
8	2027	82.1	12411	6	AX427927	Sequence
9	2025.4	82.0	9194	6	AX427926	Sequence
10	2020.4	81.8	9785	6	AX427938	Sequence
11	2018.6	81.8	9189	6	AX427921	Sequence
12	2017	81.7	9167	6	AX427933	Sequence
13	2017	81.7	9170	6	AX427928	Sequence
14	2017	81.7	9782	6	AX427935	Sequence
15	2017	81.7	9783	6	AX427934	Sequence
16	2017	81.7	9792	6	AX427932	Sequence
17	2015.6	81.6	4353	12	AF287353	Synthetic
18	1999	81.0	3015	6	AX455946	Sequence
19	1982.4	80.3	3009	6	AX455987	Sequence
20	1958.8	79.3	4352	12	AF287352	Synthetic
21	1907	77.2	2577	6	AX457088	Sequence
22	1905.8	77.2	4343	6	AX306428	Sequence
23	1901.8	77.0	4341	6	AX306429	Sequence
24	1892.2	76.6	1977	6	AX455954	Sequence
25	1879.2	76.1	7897	6	AX427923	Sequence
26	1837	74.4	1965	6	AX455952	Sequence
27	1825.6	73.9	3024	12	AF287355	Synthetic
28	1807.6	73.2	9407	6	AX427937	Sequence
29	1678.6	68.0	4307	6	AX019132	Sequence
30	1659.4	67.2	4307	6	AX035453	Sequence
31	1659.4	67.2	4307	6	AX056836	Sequence
32	1651.4	66.9	4307	6	AX283597	Sequence
33	1651.4	66.9	4307	6	AX283607	Sequence
34	1651.4	66.9	4327	6	AX035465	Sequence
35	1651.4	66.9	4353	6	AX035464	Sequence
36	1651.4	66.9	4642	6	AX035463	Sequence
37	1651.4	66.9	9772	6	AX283584	Sequence
38	1608	65.1	1680	6	AX455960	Sequence
39	1592.8	64.5	8366	6	AX180752	Sequence
40	1557.4	63.1	4338	6	AX180747	Sequence
41	1554.4	63.0	1668	6	AX455962	Sequence
42	1517.4	61.5	3981	6	AX306430	Sequence
43	1303.6	52.8	4288	6	AX149648	Sequence
44	1279.4	51.8	9010	14	AF110974	HIV-1 iso
45	1279.4	51.8	9010	14	AF110975	HIV-1 iso

ALIGNMENTS

RESULT 1	AX455914	AX455914	Sequence 30 from Patent WO0204493.	2469 bp	DNA	linear	PAT 06-JUL-2002
LOCUS	AX455914	Sequence 30 from Patent WO0204493.					
DEFINITION	AX455914	Sequence 30 from Patent WO0204493.					
ACCESSION	AX455914	Sequence 30 from Patent WO0204493.					
VERSION	AX455914.1	GI:21714899					
KEYWORDS		synthetic construct.					
SOURCE		synthetic construct					
ORGANISM		artificial sequences.					
REFERENCE		1					
AUTHORS		zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.					
TITLE		Polynucleotides encoding antigenic hiv type c polypeptides,					
JOURNAL		polypeptides and uses thereof					
		Patent: WO 0204493-A 30 17-JAN-2002;					

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

Location/Qualifiers

1..2469

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="PR975(+)"

BASE COUNT 571 a 833 c 761 g 304 t

ORIGIN

Query Match 100.0%; Score 2469; DB 6; Length 2469;

Best Local Similarity 100.0%; Pred. No. 1.3e-250;

Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGCCACCATGGCCGAGGCATGAGCCAGGCCACAGCCCAACATCTCTGATGCAG 60
DB 1 GTGAGCCACCATGGCCGAGGCATGAGCCAGGCCACAGCCCAACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGCAAGGAGGC 120
DB 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGCAAGGAGGC 120
QY 121 CATATCCCGCGAACTGCGCGGCCCGCCGCAAGAGGGCTGTGGAAAGTGGCGGAAGGAG 180
DB 121 CATATCCCGCGAACTGCGCGGCCCGCCGCAAGAGGGCTGTGGAAAGTGGCGGAAGGAG 180
QY 181 GGCCACCATGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGCCACCATGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTCCCGCAGGCAAGGCCCGCGAGTTCCCGCAGGAGAGCAACCGCGCAACAGCCCAAC 300
DB 241 TTCCCGCAGGCAAGGCCCGCGAGTTCCCGCAGGAGAGCAACCGCGCAACAGCCCAAC 300
QY 301 AGCGCGAGCTGAGGTGCGCGGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGAG 360
DB 301 AGCGCGAGCTGAGGTGCGCGGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGAG 360
QY 361 GGCAACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTTGGTGAAGATCAAGGTG 420
DB 361 GGCAACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTTGGTGAAGATCAAGGTG 420
QY 421 GGCGGCGAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGAG 480
DB 421 GGCGGCGAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGAG 480
QY 481 ATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCGCTTCAATCAAG 540
DB 481 ATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGGCGCTTCAATCAAG 540
QY 541 GTGCGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCGCATGGCACCGGTG 600
DB 541 GTGCGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCGCATGGCACCGGTG 600
QY 601 CTGATCGGCGCCCGCGGTAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCGCCCGCGGTAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAACCTTCCCGCATCAGCGCCCATCGAGACCTGCGCGTGAAGCTGAAGCCCGGATG 720
DB 661 ACCCTGAACCTTCCCGCATCAGCGCCCATCGAGACCTGCGCGTGAAGCTGAAGCCCGGATG 720
QY 721 GACGCGCCCAAGGTGAAGCAAGTGGCCCGCTGACCGAGGAGAGATCAAGGCGCCCTGACCGCC 780
DB 721 GACGCGCCCAAGGTGAAGCAAGTGGCCCGCTGACCGAGGAGAGATCAAGGCGCCCTGACCGCC 780
QY 781 ATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAAGATCGCGCCCGCGCAACCCCTAC 840
DB 781 ATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAAGATCGCGCCCGCGCAACCCCTAC 840
QY 841 AACACCCCGGTTCGCGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCGGTTCGCGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGGAC 900

QY 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCAC 960
DB 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCAC 960
QY 961 CC CGCGCGCTGAAGAAGAAGAGAGCGTACCGTGTGAGCTGGCGAGCGCTACTTC 1020
DB 961 CC CGCGCGCTGAAGAAGAAGAGAGCGTACCGTGTGAGCTGGCGAGCGCTACTTC 1020
QY 1021 AGCGTCCCTGGAGGAGGACTTCCGCAAGTACACCCCTTCCACATCCCCAGCATCAAC 1080
DB 1021 AGCGTCCCTGGAGGAGGACTTCCGCAAGTACACCCCTTCCACATCCCCAGCATCAAC 1080
QY 1081 AACGAGACCCCGCGCATCCCGCTACAGTACAAAGTGTGCCCCAGGCTGGAAGGCGAG 1140
DB 1081 AACGAGACCCCGCGCATCCCGCTACAGTACAAAGTGTGCCCCAGGCTGGAAGGCGAG 1140
QY 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCGCAACCCC 1200
DB 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCGCAACCCC 1200
QY 1201 GAGATCGTGATCTACCAAGTACATGGAGGACTGTACCTGGGCGAGCCTGGAGATCGGC 1260
DB 1201 GAGATCGTGATCTACCAAGTACATGGAGGACTGTACCTGGGCGAGCCTGGAGATCGGC 1260
QY 1261 CAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCACCCTGTGCGCTGGGGCTTCAACACC 1320
DB 1261 CAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCACCCTGTGCGCTGGGGCTTCAACACC 1320
QY 1321 CC CGACAAGAAGCACCAAGAGGAGGCCCTTCTGTGGATGGGCTAGGAGCTGCACCCC 1380
DB 1321 CC CGACAAGAAGCACCAAGAGGAGGCCCTTCTGTGGATGGGCTAGGAGCTGCACCCC 1380
QY 1381 GACAAGTGGACCGGTGACGCCATCGAGTGCCTCCGAGAGGAGAGCTGGACCGTGAACGAC 1440
DB 1381 GACAAGTGGACCGGTGACGCCATCGAGTGCCTCCGAGAGGAGAGCTGGACCGTGAACGAC 1440
QY 1441 ATCCAGAGCTGTGGGCAAGCTGAATGGGCGAGCAGATCTACCCGGGATCAAGGTG 1500
DB 1441 ATCCAGAGCTGTGGGCAAGCTGAATGGGCGAGCAGATCTACCCGGGATCAAGGTG 1500
QY 1501 CGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGCCCTGACCGACATCGTCCCTGACC 1560
DB 1501 CGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGCCCTGACCGACATCGTCCCTGACC 1560
QY 1561 GAGGAGCGCAGCTGAGCTGGCGGAGAACCGCGAGATCTGCGGAGCGCGCTGCACGGC 1620
DB 1561 GAGGAGCGCAGCTGAGCTGGCGGAGAACCGCGAGATCTGCGGAGCGCGCTGCACGGC 1620
QY 1621 GTGTACTACGACCCCGCAGGAGCCTGGTGGCGGAGATCCAGAGCAGGCGCCAGCAG 1680
DB 1621 GTGTACTACGACCCCGCAGGAGCCTGGTGGCGGAGATCCAGAGCAGGCGCCAGCAG 1680
QY 1681 TGGACCTACCAAGATCTTACGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1740
DB 1681 TGGACCTACCAAGATCTTACGAGGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAG 1740
QY 1741 ATGCGCACCGCCCGCAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCC 1800
DB 1741 ATGCGCACCGCCCGCAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCC 1800
QY 1801 ATGAGAGCATCTGTGATCTGGGGCAAGACCCCGCAAGTTCCGCGCTGCCATCCAGAGGAG 1860
DB 1801 ATGAGAGCATCTGTGATCTGGGGCAAGACCCCGCAAGTTCCGCGCTGCCATCCAGAGGAG 1860
QY 1861 ACCTGGGAGACCTGGTGGACCGACTACTGGCAGCGCAGCTGGATCCCCGAGTGGAGTTC 1920
DB 1861 ACCTGGGAGACCTGGTGGACCGACTACTGGCAGCGCAGCTGGATCCCCGAGTGGAGTTC 1920
QY 1921 GTGAACACCCCGCTGGTGAAGCTGTGTGATACAGCTGGAGAGGAGGCCATCATCGGC 1980
DB 1921 GTGAACACCCCGCTGGTGAAGCTGTGTGATACAGCTGGAGAGGAGGCCATCATCGGC 1980
QY 1981 CCCGAGACCTTCTAGGTGGACGGCGCCCAACCCCGAGACCAAGATCGGCAAGGCCCGC 2040


```
Db 1981 GCCGAGACCTTTCTAGTGGAGCGGCCGCCCAACCCGAGAGCAAGATCGGCAAGGCCGCG 2040
Qy 2041 TACGTGACCGACCGCGCGCGCAGAAAGATCGTGAGCCTTGACCGACACCAACCAAGAG 2100
Db 2041 TACGTGACCGACCGCGCGCGCAGAAAGATCGTGAGCCTTGACCGACACCAACCAAGAG 2100
Qy 2101 ACCGAGCTGACAGGCCATCCAGCTGSCCCTGCAGGACAGCGGAGGAGTGAACATCGTG 2160
Db 2101 ACCGAGCTGACAGGCCATCCAGCTGSCCCTGCAGGACAGCGGAGGAGTGAACATCGTG 2160
Qy 2161 ACCGACACCCAGTACGCCCTGGGATCATCATCAGGCCCAAGCCGACAGAGGAGCGAG 2220
Db 2161 ACCGACACCCAGTACGCCCTGGGATCATCATCAGGCCCAAGCCGACAGAGGAGCGAG 2220
Qy 2221 CTGTTGAACACAGATCATCGACGAGCTGATCAAGAGGAGAGGAGTGTACCTGAGCTGGGTG 2280
Db 2221 CTGTTGAACACAGATCATCGACGAGCTGATCAAGAGGAGAGGAGTGTACCTGAGCTGGGTG 2280
Qy 2281 CCGGCCCAACAGGGGATCGCGGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCATC 2340
Db 2281 CCGGCCCAACAGGGGATCGCGGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCATC 2340
Qy 2341 CGCAAGTGTCTTCTTCCGAGCGGATCATCGCGGCATCGTGATCTACCATGATGAC 2400
Db 2341 CGCAAGTGTCTTCTTCCGAGCGGATCATCGCGGCATCGTGATCTACCATGATGAC 2400
Qy 2401 GACCTGTACGTGGCGAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCAC 2460
Db 2401 GACCTGTACGTGGCGAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCAC 2460
Qy 2461 GGTGAATTC 2469
Db 2461 GGTGAATTC 2469

RESULT 2
AX455915
LOCUS AX455915 2463 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 31 from Patent WO020493.
ACCESSION AX455915
VERSION AX455915.1 GI:21714900
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1 zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
Patent: WO 020493-A 31 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Source
1. 2463
/organism="synthetic construct"
/db.xref="taxon:32630"
/note="PR975YM"
BASE COUNT 567 a 835 c 759 g 302 t
ORIGIN
Query Match 98.9%; Score 2442.2; DB 6; Length 2463;
Best Local Similarity 99.6%; Pred. No. 8.3e-248;
Matches 2460; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
Qy 1 GTCGACGCCACATGGCCGAGGCCATGAGCCAGCGCCACAGCGCAACATCCTGATGCAG 60
Db 1 GTCGACGCCACATGGCCGAGGCCATGAGCCAGCGCCACAGCGCAACATCCTGATGCAG 60
Qy 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGCG 120
Db 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGCG 120
Qy 121 CACATCGCCCGCACTGCGCGGCCCGCCGCAAGAGGGTCTGCTGAAGTGGCGCAAGGAG 180
```

```
Db 121 CACATCGCCCGCACTGCGCGGCCCGCCGCAAGAGGGCTGCTGAAGTGGCGCAAGGAG 180
Qy 181 GGCACACAGATGAAGAGTATGACCGAGCGCCAGGCCAACTTCTCCGGAGAGACCTGGCC 240
Db 181 GGCACACAGATGAAGAGTATGACCGAGCGCCAGGCCAACTTCTCCGGAGAGACCTGGCC 240
Qy 241 TTCCCGCCAGGGCAAGCGCGCGAGTTCCCGACGAGCAGACCCGCGCCCAACAGCCCGCC 300
Db 241 TTCCCGCCAGGGCAAGCGCGCGAGTTCCCGACGAGCAGACCCGCGCCCAACAGCCCGCC 300
Qy 301 AGCCGCGAGCTGACAGTGGCGCGGACAAACCCCGCAGCGAGCGCGCGCCAGAGCGCCAG 360
Db 301 AGCCGCGAGCTGACAGTGGCGCGGACAAACCCCGCAGCGAGCGCGCGCCAGAGCGCCAG 360
Qy 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGGAGCGCCCTGTGTGAGCATCAAGGTG 420
Db 361 GGCACCTTGAACCTTCCCGCAGATCACCTGTGGAGCGCCCTGTGTGAGCATCAAGGTG 420
Qy 421 GCGCGCCAGATCAAGGAGGCCCTGTGGACACCGCGCGGACACACCGTGTGGAGGAG 480
Db 421 GCGCGCCAGATCAAGGAGGCCCTGTGTGGACACCGCGCGGACACACCGTGTGGAGGAG 480
Qy 481 ATGAGCCTGCCCCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCAAG 540
Db 481 ATGAGCCTGCCCCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCAAG 540
Qy 541 GTGCGCCAGTACGACACAGATCCTGTAGAGATCTCGCGCAAGAGCCATCGGCAAGGAG 600
Db 541 GTGCGCCAGTACGACACAGATCCTGTAGAGATCTCGCGCAAGAGCCATCGGCAAGGAG 600
Qy 601 CTGATCGCGCCACCCCGCTGAACATCATCGCGCCCAACATCTGACCCAGCTGGGCTGC 660
Db 601 CTGATCGCGCCACCCCGCTGAACATCATCGCGCCCAACATCTGACCCAGCTGGGCTGC 660
Qy 661 ACCCTGAACCTTCCCGCATCGACCCCATCGAGACCGCTGCCCGTGAAGCTGAAGCCCGCATG 720
Db 661 ACCCTGAACCTTCCCGCATCGACCCCATCGAGACCGCTGCCCGTGAAGCTGAAGCCCGCATG 720
Qy 721 GACGCCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCGCC 780
Db 721 GACGCCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCGCC 780
Qy 781 ATCTGCGAGGAGTGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCAGAACCCCTAC 840
Db 781 ATCTGCGAGGAGTGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCAGAACCCCTAC 840
Qy 841 AACACCCCGCTGTTCGCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCGCTGTTCGCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
Qy 901 TTCCGCGAGCTGAACAAGCGCACCGACACTTCTGGGAGGTGCAGCTGGGCATCCCGCCAC 960
Db 901 TTCCGCGAGCTGAACAAGCGCACCGACACTTCTGGGAGGTGCAGCTGGGCATCCCGCCAC 960
Qy 961 CCGCGCGCCCTGAAGAAGAGAGCGTGAACCGCTGCTGGAGTGGCGAGCGCTACTTTC 1020
Db 961 CCGCGCGCCCTGAAGAAGAGAGCGTGAACCGCTGCTGGAGTGGCGAGCGCTACTTTC 1020
Qy 1021 AGCGTGCCTTGGAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAAC 1080
Db 1021 AGCGTGCCTTGGAGGAGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAAC 1080
Qy 1081 AACGAGACCCCGCATCCGCTACAGTACAAAGTGTGCCCCAGGGTGGAGGGGAGC 1140
Db 1081 AACGAGACCCCGCATCCGCTACAGTACAAAGTGTGCCCCAGGGTGGAGGGGAGC 1140
Qy 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCGCAACCCC 1200
Db 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCGCAACCCC 1200
Qy 1201 GAGATCGTGTACTACCATGATGAGCACTGTACGTGGGAGGAGCCTTGAGATCGGC 1260
Db 1201 GAGATCGTGTACTACCATGATGAGCACTGTACGTGGGAGGAGCCTTGAGATCGGC 1260
```

Db 1201 GAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGAGCAGCCTGGAGATCGGC 1254
QY 1261 CAGCACCGCGCAAGATCGAGAGCTGCGCAAGCACTCTGCTGCGTGGGGCTTCACCAACC 1320
Db 1255 CAGCACCGCGCAAGATCGAGAGCTGCGCAAGCACTCTGCTGCGTGGGGCTTCACCAACC 1314
QY 1321 CCGACAAGAAGCACCAGAGAGCCCTTCTCTGTGATGGGCTACAGCTGCACCC 1380
Db 1315 CCGACAAGAAGCACCAGAGAGCCCTTCTCTGTGATGGGCTACAGCTGCACCC 1374
QY 1381 GACAAGTGAGCGTGCAGCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACGAC 1440
Db 1375 GACAAGTGAGCGTGCAGCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACGAC 1434
QY 1441 ATCCAGAAGCTGTGGCAAGCTGAACCTGGCGCAGCAGATCTACCCGCGCATCAAGGTG 1500
Db 1435 ATCCAGAAGCTGTGGCAAGCTGAACCTGGCGCAGCAGATCTACCCGCGCATCAAGGTG 1494
QY 1501 CGCCAGCTGTCAAGCTGTGCGCGCGCCAGGCCCTGACCGACATCTGTCGCCCTGACC 1560
Db 1495 CGCCAGCTGTCAAGCTGTGCGCGCGCCAGGCCCTGACCGACATCTGTCGCCCTGACC 1554
QY 1561 GAGGAGCGAGCTGAGCTGGCGGAGAACCGGAGATCTTGGCGAGCCCTGTCACGGC 1620
Db 1555 GAGGAGCGAGCTGAGCTGGCGGAGAACCGGAGATCTTGGCGAGCCCTGTCACGGC 1614
QY 1621 GTCTACTAGCACCCAGCAAGGACCTGTGGCGGAGATCCAGAACCTGAAGTACGCGCAAG 1680
Db 1615 GTCTACTAGCACCCAGCAAGGACCTGTGGCGGAGATCCAGAACCTGAAGTACGCGCAAG 1674
QY 1681 TGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAGTACGCGCAAG 1740
Db 1675 TGGACCTACCAAGTCTACCAAGGAGCCCTTCAAGAACCTGAAGTACGCGCAAG 1734
QY 1741 ATCCGACCGCCACACCAAGGACCTGAAGCAGCTACCGAGGCCCTGAGAAAGATCGCC 1800
Db 1735 ATCCGACCGCCACACCAAGGACCTGAAGCAGCTACCGAGGCCCTGAGAAAGATCGCC 1794
QY 1801 ATGAGAGATCTGTATCTGGGCAAGACCCCAAGCTTCGCGCTGCCCATCCAGAGGAG 1860
Db 1795 ATGAGAGATCTGTATCTGGGCAAGACCCCAAGCTTCGCGCTGCCCATCCAGAGGAG 1854
QY 1861 ACCTGGAGACCTGTGTGGACCGACTACTGGCAGGCCACCTTGGATCCCCAGTGGGAGTTC 1920
Db 1855 ACCTGGAGACCTGTGTGGACCGACTACTGGCAGGCCACCTTGGATCCCCAGTGGGAGTTC 1914
QY 1921 GTCAACACCCCTGCTGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGC 1980
Db 1915 GTCAACACCCCTGCTGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGC 1974
QY 1981 GCGAGACCTTCTACGTGGACGGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGGC 2040
Db 1975 GCGAGACCTTCTACGTGGACGGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGGC 2034
QY 2041 TAGCTGACCGACCGGGCGCGGAGAGATGTGTGAGCTTACCGAGACCAACCAAGAG 2100
Db 2035 TAGCTGACCGACCGGGCGCGGAGAGATGTGTGAGCTTACCGAGACCAACCAAGAG 2094
QY 2101 ACCGAGCTGACGCGCATCCAGCTGGCCCTGACAGACAGCGGCGAGCGTGAACATCGTG 2160
Db 2095 ACCGAGCTGACGCGCATCCAGCTGGCCCTGACAGACAGCGGCGAGCGTGAACATCGTG 2154
QY 2161 ACCGAGCGAGTACGCTTGGGCATCATCCAGCGCCAGCCCGACAGAGCGAGCGAG 2220
Db 2155 ACCGAGCGAGTACGCTTGGGCATCATCCAGCGCCAGCCCGACAGAGCGAGCGAG 2214
QY 2221 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGAGAGGTGTACCTGAGCTGGGTG 2280
Db 2215 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGAGAGGTGTACCTGAGCTGGGTG 2274
QY 2281 CCGCGCCACAGGCGATCGCGGCAAGGACGATGCACAGCTGGTGAAGGCGATC 2340
Db 2275 CCGCGCCACAGGCGATCGCGGCAAGGACGATGCACAGCTGGTGAAGGCGATC 2334

QY 2341 CGCAAGTGTCTTCTTGGACGCGCATCGATGGCGGATCGTGTATCTACCAAGTACATGGAC 2400
Db 2335 CGCAAGTGTCTTCTTGGACGCGCATCGATGGCGGATCGTGTATCTACCAAGTACATGGAC 2394
QY 2401 GACCTGTAGTGGGAGCGGCGCCCTAGATCGATTAAAGCTTCCCGGGGTAGCACC 2460
Db 2395 GACCTGTAGTGGGAGCGGCGCCCTAGATCGATTAAAGCTTCCCGGGGTAGCACC 2454
QY 2461 GGTGAATTC 2469
Db 2455 GGTGAATTC 2463

RESULT 3
AX455916
LOCUS AX455916 Sequence 32 from Patent WO0204493. linear PAT 06-JUL-2002
DEFINITION AX455916
ACCESSION AX455916
VERSION AX455916.1 GI:21714901
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Polypeptides and uses thereof
PUBLISHED Patent: WO 0204493-A 32 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
1..2457
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975YMHM"

BASE COUNT 566 a 837 c 754 g 300 t
ORIGIN

Query Match 97.8%; Score 2415.4; DB 6; Length 2457;
Best Local Similarity 99.3%; Pred. No. 5.4e-245;
Matches 2451; Conservative 0; Mismatches 6; Indels 12; Gaps 2;

QY 1 GTCGAGCCACCATGCGCGAGGCCATGAGCCAGGCCACAGGCCACACATCTCTGATGCAG 60
Db 1 GTCGAGCCACCATGCGCGAGGCCATGAGCCAGGCCACAGGCCACACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTTCAACTGCGGCAAGGAGGC 120
Db 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTTCAACTGCGGCAAGGAGGC 120
QY 121 CACATGCGCGCAACTGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGGAGAG 180
Db 121 CACATGCGCGCAACTGCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGGAGAG 180
QY 181 GGCACACAGATGAAGGACTGCACCGGAGGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
Db 181 GGCACACAGATGAAGGACTGCACCGGAGGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTCCCCCGAGGCAAGCCCGCGAGTTCCCCCAGCGAGCAACCGCCCAACAGCCCCCACC 300
Db 241 TTCCCCCGAGGCAAGCCCGCGAGTTCCCCCAGCGAGCAACCGCCCAACAGCCCCCACC 300
QY 301 AGCGCGAGCTGCAAGTGTGCGGGCGACACCCCGCAGCGAGGCCGCGCGCGAGCGCAG 360
Db 301 AGCGCGAGCTGCAAGTGTGCGGGCGACACCCCGCAGCGAGGCCGCGCGCGAGCGCAG 360
QY 361 GGCACCCCTGAACCTTCCCGCAGATCACCCCTGTGGAGCGCCCTGTGTGACATCAAGGTG 420
Db 361 GGCACCCCTGAACCTTCCCGCAGATCACCCCTGTGGAGCGCCCTGTGTGACATCAAGGTG 420
QY 421 GCGGGCCAGATCAAGGAGGCCCTGTGGACACCGCGCCCGCAGCACACCGTGTGGAGGAG 480
Db 421 GCGGGCCAGATCAAGGAGGCCCTGTGGACACCGCGCCCGCAGCACACCGTGTGGAGGAG 480

QY	481	ATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGTCGGCGGCATCGSGCGCTTCATCAAG	540
Db	481		
Db	481	ATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGCATCGSGCGCTTCATCAAG	540
QY	541	GTCGGCAGTAGACACAGATCCTGTATCGAGATCTTGCGGCAAGAAGGCCATCGGCAACGGTG	600
Db	541		
Db	541	GTCGGCAGTAGACACAGATCCTGTATCGAGATCTTGCGGCAAGAAGGCCATCGGCAACGGTG	600
QY	601	CTGATCGGCCCCACCCTCGGTGAACATATCGGCGCGCAACATGCTGACCCAGCTGGGCTGC	660
Db	601		
Db	601	CTGATCGGCCCCACCCTCGGTGAACATATCGGCGCGCACATGCTGACCCAGCTGGGCTGC	660
QY	661	ACCTTGAATCTTCCCATCAGCCCCATCGAGACCGTGGCCGTGTAAGCTGAAAGCCCCGGCATG	720
Db	661		
Db	661	ACCTTGAATCTTCCCATCAGCCCCATCGAGACCGTGGCCGTGTAAGCTGTAAGCGCGGCATG	720
QY	721	GACGGCCCCAAGTGAAGCAGTGGCCCTGACCCGAGGAGAGATCAAGGCCCTGACCGCC	780
Db	721		
Db	721	GACGGCCCCAAGTGAAGCAGTGGCCCTGACCCGAGGAGAGATCAAGGCCCTGACCGCC	780
QY	781	ATCTGGAGGAGATGSAGAAGGAGGGCAAGATCAACRAAGATCGGCCCGGAGAACCCCTAC	840
Db	781		
Db	781	ATCTGCGAGGAGATGSAGAAGGAGGGCAAGATCAACRAAGATCGGCCCGGAGAACCCCTAC	840
QY	841	AACACCCCGTGTTCGCCATCAAGAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGAGC	900
Db	841		
Db	841	AACACCCCGTGTTCGCCATCAAGAAGAAAGAGACACCAAGTGGCGCAAGCTGGTGAGC	900
QY	901	TTCGGGAGCTGAACAAGCGCACCCAGAGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC	960
Db	901		
Db	901	TTCGGGAGCTGAACAAGCGCACCCAGAGACTTCTGGGAGGTGCAGCTGGGCATCCCCCAC	960
QY	961	CCGCGCGGCTGAAGAAGAAAGAGCGGTGACCGTGTGGACGTGGGCGACGCTACTTC	1020
Db	961		
Db	961	CCGCGCGGCTGAAGAAGAAAGAGCGGTGACCGTGTGGACGTGGGCGACGCTACTTC	1020
QY	1021	AGCGTGGCCCTGGACGAGACATTCCGCAAGTACACCGCTTTCACCATCCCCAGCATCAAC	1080
Db	1021		
Db	1021	AGCGTGGCCCTGGACGAGACATTCCGCAAGTACACCGCTTTCACCATCCCCAGCATCAAC	1080
QY	1081	AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC	1140
Db	1081		
Db	1081	AACGAGACCCCGGCATCCGCTACCAGTACAACGTGCTGCCCCAGGGCTGGAAGGGCAGC	1140
QY	1141	CCGAGCATCTTCAGAGCAGCATGACCAAGATCTCTGAGAGCCCTTCCGCGCCCGCAACCCC	1200
Db	1141		
Db	1141	CCGAGCATCTTCAGAGCAGCATGACCAAGATCTCTGAGAGCCCTTCCGCGCCCGCAACCCC	1200
QY	1201	GAGATCGTGATCTACCAAGTACATGACGACGCTGTACGTGGCGACGACCTGGAGATCGGC	1260
Db	1201		
Db	1201	GAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGACGACCTGGAGATCGGC	1260
QY	1261	CAGCACCGGCCAAGATCGAGGAGCTGGGCAAGCACCTGCTGCCTGGGGCTTCACCAACC	1320
Db	1261		
Db	1261	CAGCACCGGCCAAGATCGAGGAGCTGGGCAAGCACCTGCTGCCTGGGGCTTCACCAACC	1320
QY	1321	CCGAGACAAAGACCCAGAGAGCCCTTCCCTGTGGATGGGCTAGGAGCTGCACCCC	1380
Db	1321		
Db	1321	CCGAGACAAAGACCCAGAGAGCCCTTCCCTGTGGATGGGCTAGGAGCTGCACCCC	1380
QY	1381	GACAAGTGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGACTGGACCGTGAAGAC	1440
Db	1381		
Db	1381	GACAAGTGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGACTGGACCGTGAAGAC	1440
QY	1441	ATCCAGAAGCTGTGGGCAAGCTGAACTGGGCCACCCAGATCTACCCCGGCATCAAGGTG	1500
Db	1441		
Db	1441	ATCCAGAAGCTGTGGGCAAGCTGAACTGGGCCACCCAGATCTACCCCGGCATCAAGGTG	1500
QY	1501	CGCCAGCTGTGCAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGACATCGTGCCCTGACC	1560
Db	1501		
Db	1501	CGCCAGCTGTGCAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGACATCGTGCCCTGACC	1560
QY	1548	CGCAGCTGTGCAAGCTGTGGCGGCGCCAAAGGCCCTGACCGACATCGTGCCCTGACC	1608
Db	1548		
Db	1548	CGCAGCTGTGCAAGCTGTGGCGGCGCCAAAGGCCCTGACCGACATCGTGCCCTGACC	1608

ORGANISM	synthetic construct		
REFERENCE	artificial sequences.		
AUTHORS	1		
TITLE	Huang, Y. and Nabel, G.J.		
JOURNAL	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization		
FEATURES	Patent: WO 0232943-A 168 25-APR-2002;		
Source	GOVERNMENT OF THE UNITED STATES (US)		
BASE COUNT	Location/Qualifiers		
ORIGIN	1..9166		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="plasmid pVR1012x/s containing HIV genes"		
	2227 a 2662 c 2513 g 1764 t		
Query Match	83.1%; Score 2051.2; DB 6; Length 9166;		
Best Local Similarity	92.6%; Pred. No. 6.6e-207;		
Matches 2192; Conservative	0; Mismatches 163; Indels 13; Gaps 3;		
QY 14	TGGCCGAGCCATGAGCCAGGCGACAGCCCAACATCTGTGTCAGCGCGCAACTTCA	73	
DB 2961	TGGCCGAGCCATGAGCAAGGTGAACACCAACATCATGATGTCAGCGCAACTGCA	3020	
QY 74	AGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACATGCGCCGCA	133	
DB 3021	AGGGCCCCAAGCGCACCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACCTGGCCCGCA	3080	
QY 134	ACTGCGCGCGCCCGCCAGAGAGGGTGTGAAAGTGGCGGCAAGGAGGCCACAGATGA	193	
DB 3081	ACTGCGCGCGCCCGCCAGAGAGGGTGTGAAAGTGGCGGCAAGGAGGCCACAGGTGA	3140	
QY 194	AGGACTGCAACCGAGCGCCAGGCAACTTCTCGCGAGGACCTGGCTTCCCGAGGGCA	253	
DB 3141	AGGACTGCAACCGAGCGCACAGGTAA-TTTTTAGGAAGATCTGGCTTCCACAGGGA	3199	
QY 254	AGGCGCGGAGTTCCCGAGCGAGAGAACCGCGCAACAGCCCGACAGCGCGAGCTGC	313	
DB 3200	AGGCGAGGAAATTTCTTTCAGAGCAGACAGAGCCAAACAGCCCGACAGAGAGCTTC	3259	
QY 314	AGGTGGCGGG-----CGACACCCCGGAGGAGCGCGCGCGCGAGCGCGAGGGA---	364	
DB 3260	AGGTGGGGAGAGACAACAACCTCCCTCAGAGCAGGAGCGCGCGCGCGATAGACAAGAACTG	3319	
QY 365	----CCTGAACCTTCCCGAGATCACCTGTGCGAGCGCCCTGTGTGACATCAAGGTGG	421	
DB 3320	TATCTTTAGCTTCCCTCAGATCACTCTTTGCGAGCGACCCCTCGTCAAAATAGATAG	3379	
QY 422	GGGGCAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGCACACCGTGTGGAGGAGA	481	
DB 3380	GGGGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGAGCACACCGTGTGGAGGAGA	3439	
QY 482	TGAGCTGCGCGCGAGTGAAGCCCAAGATGATCGGGCGCATCGGGCGCTTCATCAAGG	541	
DB 3440	TGAACCTGCGCGCGCTGGAGCGCCCAAGATGATCGGGCGCATCGGGCGCTTCATCAAGG	3499	
QY 542	TGCGCGAGTACACAGATCTGTGAGATCTGTGGGCAAGAGGCCATCGGCACCGTGC	601	
DB 3500	TGGCCAGTACACAGATCTGTGAGATCTGTGGGCAAGAGGCCATCGGCACCGTGC	3559	
QY 602	TGATGCGCGCGCGGTGAACATCATGCGCGCGCAACATGTCACCGAGCTGGCTGCA	661	
DB 3560	TGTTGGCGCGCGCGGTGAACATCATCGCGCGCAACATGTCACCGAGCTGGCTGCA	3619	
QY 662	CCCTGAACCTTCCCGATCAGCCCGCATCGAGACCGTGCCTGTGAGCTGAGCGCGCATGG	721	
DB 3620	CCCTGAACCTTCCCGATCAGCCCGCATCGAGACCGTGCCTGTGAGCTGAGCGCGCATGG	3679	
QY 722	ACGGCCCCAAGGTGAAGAGTGGCGCTTGAACCGAGGAGAGATCAAGGCGCTTGAACCGCA	781	
DB 3680	ACGGCCCCAAGGTGAAGAGTGGCGCTTGAACCGAGGAGAGATCAAGGCGCTTGAACCGCA	7339	
QY 782	TCGCGAGGAGATGGAAGAGGAGGCAAGATCAACCAAGATCGGCGCGCGAGAACCCCTACA	841	
DB 3740	TCGCGAGGAGATGGAAGAGGAGGCAAGATCAACCAAGATCGGCGCGCGAGAACCCCTACA	3799	
QY 842	ACACCCCGGTGTTGCGCATCAAGAAAGGACAGCACCAAGTGGCGCAAGCTGGTGACT	901	
DB 3800	ACACCCCGGTGTTGCGCATCAAGAAAGGACAGCACCAAGTGGCGCAAGCTGGTGACT	3859	
QY 902	TGCGCGAGCTGAACAGCGCCAGCCAGGACTTCTGGAGGTGTCAGCTGGGCATCCCCACC	961	
DB 3860	TGCGCGAGCTGAACAGCGCCAGCCAGGACTTCTGGAGGTGTCAGCTGGGCATCCCCACC	3919	
QY 962	CGCCCGCTTGAAGAAAGAGAGCGTGTACCGTGTGTGACGTGGGCGAGCCCTACTTCA	1021	
DB 3920	CGCCCGCTTGAAGAAAGAGAGCGTGTACCGTGTGTGACGTGGGCGAGCCCTACTTCA	3979	
QY 1022	GGTGCGCTTGGAGAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGATCAACA	1081	
DB 3980	GGTGCGCTTGGAGAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGATCAACA	4039	
QY 1082	ACGAGACCCCGCATCCGCTACCATGACAGTACAACTGTGCTGCGCGCGCGGAGGCGCC	1141	
DB 4040	ACGAGACCCCGCATCCGCTACCATGACAGTACAACTGTGCTGCGCGCGCGGAGGCGCC	4099	
QY 1142	CCAGCATCTTCCAGAGCAGCATGACCAAGATCTGTGAGCGCTTCCGCGCGCGGAGGCGCC	1201	
DB 4100	CCAGCATCTTCCAGTACAGCATGACCAAGATCTGTGAGCGCTTCCGCGAGCAACCCCG	4159	
QY 1202	AGATCTGTATACCATGATGACAGCTGTACCTGTGGGCGAGCGCTGGAGATCGGCC	1261	
DB 4160	ACATCTGTATACCATGATGACAGCTGTACCTGTGGGCGAGCGCTGGAGATCGGCC	4219	
QY 1262	AGCACCGCGCGAGATCGAGGAGCTGCGCAAGCATCTGTGCTGGGCTTCCAGCACCC	1321	
DB 4220	AGCACCGCGCGAGATCGAGGAGCTGCGCAAGCATCTGTGCTGGGCTTCCAGCACCC	4279	
QY 1322	CGGACAAGAGCACCAGAGAGGCGCCCTTCTGTGGATGGCTTACGAGCTCAGCCCG	1381	
DB 4280	CGGACAAGAGCACCAGAGAGGCGCCCTTCTGTGGATGGCTTACGAGCTCAGCCCG	4339	
QY 1382	ACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAAGACACA	1441	
DB 4340	ACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAAGACACA	4399	
QY 1442	TCAGAGCTGTGGGCAAGCTGAAGTGGCGCGCGAGCTTACCCCGCGCATCAAGGTGC	1501	
DB 4400	TCAGAGCTGTGGGCAAGCTGAAGTGGCGCGCGAGCTTACCCCGCGCATCAAGGTGC	4459	
QY 1502	GCCAGCTGTCAAGCTGTGCGCGCGCGCAAGGCGCTTACCGACATCTGTGCCCTTACCCG	1561	
DB 4460	GCCAGCTGTCAAGCTGTGCGCGCGCGCAAGGCGCTTACCGAGGTGTGTGCCCTTACCCG	4519	
QY 1562	AGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTTCCGCGAGCGCTGGCAGCGG	1621	
DB 4520	AGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTTCCGAGGAGCGCTGGCAGCGG	4579	
QY 1622	TGTACTTACACCCAGCAAGAGCTGTGTCGCGAGATCCAGAGAGCGGCGAGTCCAGT	1681	
DB 4580	TGTACTTACACCCAGCAAGAGCTGTGTCGCGAGATCCAGAGAGCGGCGAGTCCAGT	4639	
QY 1682	GGACCTTACAGATCTACAGAGGCGCTTCAAGAACCTTGAAGACCGCGCAAGTACGCCAAGA	1741	
DB 4640	GGACCTTACAGATCTACAGAGGCGCTTCAAGAACCTTGAAGACCGCGCAAGTACGCCAAGA	4699	
QY 1742	TGCGCACCGCCCGACACAGAGCTGAGCAGCTGACCGAGGCGCTGGCAGAGATCGCCA	1801	
DB 4700	TGAAGGCGCGCGCACACCAACAGCTGAAGCAGCTGACCGAGGCGCTGGCAGAGATCGCCA	4759	
QY 1802	TGAGAGCATCTGTGATCTGGGCGCAAGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT	1861	
DB 4760	CCGAGAGCATCTGTGATCTGGGCGCAAGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT	4819	
QY 1862	CCTGGGAGACCTTGTGGAGCCGACTGTGCGAGGCGCCACTTGGATCCCGGAGTGGAGTTCCG	1921	

```
Db 4820 CCTGGAGGCGCTGGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGGAGTGGAGTTGG 4879
QY 1922 TGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGCG 1981
Db 4880 TGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGCG 4939
QY 1982 CGAGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGAGACCAAGATCGGCAAGCGCGGT 2041
Db 4940 CGAGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGAGACCAAGATCGGCAAGCGCGGT 4999
QY 2042 ACGTGAACAGCGGCGCGCAGAGAACTGTGTAGCCTGACCGAGAGACCAACCAAGA 2101
Db 5000 ACGTGAACAGCGGCGCGCAGAGAACTGTGTAGCCTGACCGAGAGACCAACCAAGA 5059
QY 2102 CGAGCTGAGGCGCATTCAGCTGGCCCTGCAGGACAGCGCAGCGAGGTGAACATCGTGA 2161
Db 5060 CGAGCTGAGGCGCATTCAGCTGGCCCTGCAGGACAGCGCCTGGAGGTGAACATCGTGA 5119
QY 2162 CGACAGCCAGTAGTACGCCCTGGGCGATCATCCAGGCCAGCCGCGCAAGAGCGAGCGAGC 2221
Db 5120 CGACAGCCAGTAGTACGCCCTGGGCGATCATCCAGGCCAGCCGCGCAAGAGCGAGCGAGC 5179
QY 2222 TGTGAACAGCATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 2281
Db 5180 TGTGAACAGCATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 5239
QY 2282 CGCCCGCAGGCGCATCGCGCGCAAGCAGCAGATCGACAAGCTGGTGAGCAAGGCGATCC 2341
Db 5240 CGCCCGCAGGCGCATCGCGCGCAAGCAGCAGATCGACAAGCTGGTGAGCGCGCGATCC 5299
QY 2342 GCAAGTGTCTTCTCGAGCGGCGATCGA 2369
Db 5300 GCAAGTGTCTTCTCGAGCGGCGATCGA 5327

RESULT 5
AX427936
LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 174 from Patent WO0232943.
ACCESSION AX427936
VERSION AX427936.1 GI:21538023
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE
1 Huang, Y. and Nabel, G.J.
AUTHORS Modifications of hiv env, gag, and pol enhance immunogenicity for
TITLE genetic immunization
JOURNAL Patent: WO 0232943-A 174 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Source Location/Qualifiers
1..9788
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2377 a 2817 c 2695 g 1899 t
ORIGIN

Query Match 82.7%; Score 2042.8; DB 6; Length 9788;
Best Local Similarity 91.8%; Pred. No. 5e-206;
Matches 2197; Conservative 0; Mismatches 177; Indels 18; Gaps 3;

QY 14 TGGCCGAGGCGCATGAGCCAGCCAGCCCAACATCTCTGATGCGAGCGCAACTTCA 73
Db 2961 TGGCCGAGGCGCATGAGCAAGGTGAACCAACCAACATCATGATGCGAGCGCAACTTCA 3020
QY 74 AGGCCCGCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCGCA 133
Db 3021 AGGCCCGCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGGCCACCTGGCCGCA 3080
QY 134 ACTGCGCGCGCCCGCAAGAGGCGTGTGGAAGTGCAGAGGAGGCGCACAGATGA 193
Db 4214 ACTGCGCGCGCCCGCAAGAGGCGTGTGGAAGTGCAGAGGAGGCGCACAGATGA 4214
```

```
Db 3081 ACTGCGCGCGCCCGCCCAAGAGAGGGCTGTGTGAAGTGCAGCAAGAGGCGCACAGGTGA 3140
QY 194 AGGACTGCACCGAGCGCCAGCGCCCAACTTCTTCGCGAGGAGCTTGGCTTCCCCAGGGCA 253
Db 3141 AGGACTGCACCGAGCGCACAGGCTAA-----TAGGGAAGATCTGGCTTCCCCAAGGGA 3194
QY 254 AGCCCGCGAGTTTCCCGAGCGAGCAGAACCGCGCCACACAGCCCGCCAGCGGAGCTGC 313
Db 3195 AGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACACGCCCCACCAGAGAGAGCTTC 3254
QY 314 AGTGTGGCGG-----CGACAACCCCGCAGAGGCGCGGCGAGCGCGCAGGGCA--- 364
Db 3255 AGTGTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAACTG 3314
QY 365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTTGGTGAGCATCAAGTGG 421
Db 3315 TATCTCTTAGCTTCCCTCAGATCCTCTTGGCAGCAGCCCTCGTGCACATAAAGATAG 3374
QY 422 GCGGCCAGATCAAGGAGGCGCTGTGGACACCGCGCGCCGACACACCGTGTCTGGAGAGA 481
Db 3375 GGGGCCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCAGACACCGTGTCTGGAGAGA 3434
QY 482 TGAGCCTGCCCGGGAAGTGAAGCCCAAGATGATCGCGCGGCGATCGCGGCTTCATCAAG 541
Db 3435 TGAACCTGCCCGCGCGTGGAAAGCCCAAGATGATCGCGCGGCGATCGCGGCTTCATCAAG 3494
QY 542 TGCGCCAGTACGACACAGATCCTGATCGAGATCTCGGCAAGAGGCCATCGGACCGTGC 601
Db 3495 TGGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCCACAGGCCATCGGACCGTGC 3584
QY 602 TGATCGCCCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCCAGTGGGCTGCA 661
Db 3555 TGGTGGSCCGCCACCGCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCTGCA 3614
QY 662 CCCTGAACCTCCCATCAGCCCGATCGAGACCGTGCCTGCGGCGGATGG 721
Db 3615 CCCTGAACCTCCCATCAGCCCGATCGAGACCGTGCCTGCGGCGGATGG 3674
QY 722 ACGCCCGCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGCGCTGACCGCCA 781
Db 3675 ACGCCCGCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGCGCTGCTGGAGA 3734
QY 782 TCTCGAGGAGATGGAGAAGGCGGCAAGATCAACAGATCGGCCCGCGAGAACCCCTACA 841
Db 3735 TCTGACCGAGATGGAGAAGGCGGCAAGATCAACAGATCGGCCCGCGAGAACCCCTACA 3794
QY 842 ACACCCCGCTGTTCGCCATCAAGAGAGAGACACACCAAGTGGCGGAGCTGTGGACT 901
Db 3795 ACACCCCGCTGTTCGCCATCAAGAGAGAGACACACCAAGTGGCGGAGCTGTGGACT 3854
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACCTGGGCATCCCCCACC 961
Db 3855 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACCTGGGCATCCCCCACC 3914
QY 962 CCGCGCGCTTGAAGAAGAAGAGCGCTGACCGTGTGGAGTGGGCGGACCGCTACTTCA 1021
Db 3915 CCGCGCGCTTGAAGAAGAAGAGCGCTGACCGTGTGGAGTGGGCGGACCGCTACTTCA 3974
QY 1022 GCGTGGCCCTTGGAGGAGACTTCCGCAAGTACACCGCGCTTCACCATCCCGAGCATCAACA 1081
Db 3975 GCGTGGCCCTTGGAGGAGACTTCCGCAAGTACACCGCGCTTCACCATCCCGAGCATCAACA 4034
QY 1082 ACAGAGCCCGCGGATCCGCTACCAAGTACACCGCGCTTCACCATCCCGAGCATCAACA 1141
Db 4035 ACAGAGCCCGCGGATCCGCTACCAAGTACACCGCGCTTCACCATCCCGAGCATCAACA 4094
QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCCAGCCCG 1201
Db 4095 CCGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCCAGCCCG 4154
QY 1202 AGATCGTGTATCTACCATGATGACACCGCTGTAGTGGGAGGAGCTGGAGATCGGCC 1261
Db 4155 ACATCGTGTATCTACCATGATGACACCGCTGTAGTGGGAGGAGCTGGAGATCGGCC 4214
```

QY 1262 AGCACCGCCGACAGATCGAGGAGCTGCGAAGCAACCTGCTGCGCTGGGCTTCAACACCC 1321
|||||
Db 4215 AGCACCGCCGACAGATCGAGGAGCTGCGCAGCACTGCTGCGCTGGGCTTCAACACCC 4274
|||||
QY 1322 CCGACAAGAACACCAAGAGAGGAGCCCTTCTCTGTGGATGGCTACGAGCTGCACCCCG 1381
|||||
Db 4275 CCGACAAGAACACCAAGAGAGGAGCCCTTCTCTGTGGATGGCTACGAGCTGCACCCCG 4334
|||||
QY 1382 ACAAGTGGACCTGTCAGCCCATCGAGCTGCCCGAGAGAGAGAGCTGGACGTGAACGACA 1441
|||||
Db 4335 ACAAGTGGACCTGTCAGCCCATCGAGCTGCCCGAGAGAGAGAGCTGGACGTGAACGACA 4394
|||||
QY 1442 TCAGAAAGTGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCGCGCATCAAGGTGC 1501
|||||
Db 4395 TCAGAAAGTGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCGCGCATCAAGGTGC 4454
|||||
QY 1502 GCCAGCTGTCAAGCTGTGCGCGGCGCCCAAGCCCTGACCGACATCGTCCCTGACCG 1561
|||||
Db 4455 GCCAGCTGTCAAGCTGTGCGCGGCGCCCAAGCCCTGACCGAGGTGTGCTCCCTGACCG 4514
|||||
QY 1562 AGGAGCCGAGCTGGAGCTGGCGGAGACCGGAGATCTGCGGAGCCGCTGCACGGCG 1621
|||||
Db 4515 AGGAGCCGAGCTGGAGCTGGCGGAGACCGGAGATCTGCGGAGCCGCTGCACGGCG 4574
|||||
QY 1622 TGTACTTACGACCCAGCAAGGACCTGTGTGGCCGAGATCCAGAGCAGGCGCCACGACCT 1681
|||||
Db 4575 TGTACTTACGACCCAGCAAGGACCTGTGTGGCCGAGATCCAGAGCAGGCGCCACGACCT 4634
|||||
QY 1682 GGACCTACAGATCTACCAAGAGCCCTTCAAGACCTGAAGACCGGCAAGTACGCCAAGA 1741
|||||
Db 4635 GGACCTACAGATCTACCAAGAGCCCTTCAAGACCTGAAGACCGGCAAGTACGCCAAGA 4694
|||||
QY 1742 TCGCACCCGCCACACACGACGCTGAACAGCTGACCCAGGCGCTGCAGAGATCGCCA 1801
|||||
Db 4695 TGAAGGCGCCGACACCAACAGAGCTGAAGAGCTGACCCAGGCGCTGCAGAGATCGCCA 4754
|||||
QY 1802 TGGAGAGCATCTGTGCTGGGCAAGACCCCAAGTTCGCCCTGCCATCCAGAGGAGA 1861
|||||
Db 4755 CCGAGAGCATCTGTGCTGGGCAAGACCCCAAGTTCAGCTGCCATCCAGAGGAGA 4814
|||||
QY 1862 CTTGGAGAGCTGTGGACCGACTACTGCGAGGCGCACCTGGATTCGCCGAGTGGGAGTTCG 1921
|||||
Db 4815 CTTGGAGAGCTGTGGACCGACTACTGCGAGGCGCACCTGGATTCGCCGAGTGGGAGTTCG 4874
|||||
QY 1922 TGAACACCCCGCTGTGTGAGAGCTGTGTACAGCTGGAGAGGAGCCATCATCGGCG 1981
|||||
Db 4875 TGAACACCCCGCTGTGTGAGAGCTGTGTACAGCTGGAGAGGAGCCATCATCGGCG 4934
|||||
QY 1982 CCGAGACCTTCTACGTGGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGGCT 2041
|||||
Db 4935 CCGAGACCTTCTACGTGGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGGCT 4994
|||||
QY 2042 AGCTGACGACCGGCGCGGCAAGAGATCGTGAGCTGACCGAGACCAACCAACCAAGA 2101
|||||
Db 4995 AGCTGACGACCGGCGCGGCAAGAGATCGTGAGCTGACCGAGACCAACCAACCAAGA 5054
|||||
QY 2102 CCGAGCTGACGACCATCAGCTGGCCCTGCGAGGACGAGGAGGAGTGAACATCGTGA 2161
|||||
Db 5055 CCGAGCTGACGACCATCAGCTGGCCCTGCGAGGACGAGGAGGAGTGAACATCGTGA 5114
|||||
QY 2162 CCGACGACGAGTACGCTGGGACATCTCCAGGCGCCAGCCGACAGAGGAGGAGCGAGC 2221
|||||
Db 5115 CCGACGACGAGTACGCTGGGACATCTCCAGGCGCCAGCCGACAGAGGAGGAGCGAGC 5174
|||||
QY 2222 TGGTGAACAGATCATCAGAGCTGTATCAAGAGGAGAGTGTACCTGAGCTGGGTGC 2281
|||||
Db 5175 TGGTGAACAGATCATCAGAGCTGTATCAAGAGGAGAGTGTACCTGAGCTGGGTGC 5234
|||||
QY 2282 CCGCCCAACAGGCGCATCGCGGCAACGAGCAGATCGAAGCTGTGTAGCAGAGGACATCC 2341
|||||
Db 5235 CCGCCCAACAGGCGCATCGCGGCAACGAGCAGATCGAAGCTGTGTAGCAGGCGCATCC 5294
|||||

QY 2342 GCAAGGTGCTGTCTCTGGAGCGCATCGATGGCGGCATCGTGTATCTACCACTA 2393
|||||
Db 5295 GCAAGGTGCTGTCTCTGGAGCGCATCGATGGCGGCATCGACAAAGCCCGAGGAGCAGAGAA 5346
|||||
RESULT 6
AX427931
LOCUS AX427931 9169 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 169 from Patent WO0232943.
ACCESSION AX427931
VERSION AX427931.1 GI:21538018
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 169 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source
1..9169
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2227 a 2668 c 2505 g 1769 t
ORIGIN

Query Match 82.2%; Score 2028.8; DB 6; Length 9169;
Best Local Similarity 91.7%; Pred. No. 1.5e-204;
Matches 2194; Conservative 0; Mismatches 182; Indels 16; Gaps 4;

QY 14 TGGCCGAGCCCATGAGCCAGGCCACCGCCCAACATCTCTGATGCGAGCGCAGCAATTCA 73
|||||
Db 2967 TGGCCGAGCCCATGAGCCAGGCCACCGCCCAACATCTCTGATGCGAGCGCAGCAATTCA 3026
|||||
QY 74 AGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCCCGCA 133
|||||
Db 3027 AGGCCCCCAAGCGCAT--CAAGTCTTCACTGCGGCAAGGAGGCGCCACCTGGCCCGCA 3083
|||||
QY 134 ACTGCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGCAGGAGGCGCCACAGATGA 193
|||||
Db 3084 ACTGCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGCAGGAGGCGCCACAGATGA 3143
|||||
QY 194 AGGACTGCGAGCGCCAGCGCAACTTCTTCCGCGAGGACCTGGCTTCCCGCCAGGCGCA 253
|||||
Db 3144 AGGACTGCGAGCGCGAGCGACAGGCTAA-TTTTATTAGGAAGATCTGGCTTCCCGCAAGGGA 3202
|||||
QY 254 AGGCCCCGAGTTCCCGAGCGAGCAACCGCGCCCAACAGCCCGCCAGCGCGAGCTGC 313
|||||
Db 3203 AGGCCCCGAGTTTCTTCAGAGCAGCAGAGCCAGAGCCCGCCAGCAGAGAGAGCTTC 3262
|||||
QY 314 AGGTGCGCGG-----CGACAACCCCGCAGGAGGCGCGCGCCAGCGCCAGGCGCA--- 364
|||||
Db 3263 AGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAGCAGGAGCGCGATAGAGAAGAACTG 3322
|||||
QY 365 ---CCCTGACATTTCCCGAGATCACCTGTGGCAGGCGCCCTTGGTGAGCATCAAGGTGG 421
|||||
Db 3323 TATCTTTTATGCTTCCCTCAGATCACTCTTTTGGCAGCGACGCTCGCTCACAATAAAGATAG 3382
|||||
QY 422 GGGCCAGATCAAGAGGCGCTGTGGACACCGCGCGCCGACGACACCGCTGCTGGAGGAGA 481
|||||
Db 3383 GGGCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCCGACGACACCGCTGCTGGAGGAGA 3442
|||||
QY 482 TGAGCTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGATCGGCGGCTTTCATCAAGG 541
|||||
Db 3443 TGAACCTGCCCGCGCTGGAAGCCCAAGATGATCGGCGGATCGGCGGCTTTCATCAAGG 3502
|||||
QY 542 TCGCGCAGTACACAGATCTCTGATCGAGATCTGGGCGAAGAGGCGCCATCGGACCGTGC 601
|||||
Db 3503 TGGCGCATACGACAGATCTCTGATCGAGATCTGGCGGCCAACAGGCCATCGGACCGTGC 3562
|||||

Best Local Similarity 91.6%; Pred. No. 2.3e-204; Matches 2194; Conservative 0; Mismatches 185; Indels 16; Gaps 4;			
QY	14	TGGCCGAGGCCATGAGCCAGG---CCACCAGCGCAACATCTGATGAGCGCAGCAACT	70
Db	2967	TGGCCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATATGATGAGCGCGCAACT	3026
QY	71	TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGAGGCCACATCGGCC	130
Db	3027	TCCGCAACCGCAAGATCGTGAAGTCTTCAACTGCGCAAGAGAGGCCACACCGCCC	3086
QY	131	GCAACTGCCGGCCCCCGGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACCAGA	190
Db	3087	GCAACTGCCGGCCCCCGGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACCAGA	3146
QY	191	TGAAGGACTGCACGAGGCGCCAGCCAACTTCTTCCCGAGGAGCTGGCTTCCGCCAGG	250
Db	3147	TGAAGGACTGCACGAGGCGCACAGGCTAA-TTTTTTAGGGAAGATCTGGCTTTCCACAG	3205
QY	251	GCAAGGCGCGAGTTTCCCGAGCAGAGACCGCGCAACAGAGCCCCACCAGCGCAGC	310
Db	3206	GGAAGGCCAGGCAATTTTCTTCAGAGCAGACAGAGCCAAACAGCCCCACAGAAGAGC	3265
QY	311	TGCAGGTGCGGG-----CGACAACCCCCGAGCGAGGCGCGGCCGAGCGCCAGGSCA	364
Db	3266	TTCAAGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAAAGCAGGAGCCGATAGACAAGAA	3325
QY	365	-----CCCTGAACCTTCCCCAGATCACCTGTGGCAGCGCCCTGGTGAGCATCAAG	418
Db	3326	CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCTGTCACAAATAAGA	3385
QY	419	TGGCGCGCAGATCAAGAGGCGCTGTGGACACCGGCGCGCAGACACCGTGTGGAGG	478
Db	3386	TAGGGGCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCAGACACCGTGTGGAGG	3445
QY	479	AGATGAGCTGCCCGGCAAGTGAAGCCCAAGATGATCGGGCGCATCGCGGCTTCATCA	538
Db	3446	AGATGAACCTGCCCGGCGCTTGAAGGCCCAAGATGATCGGGCGCATCGCGGCTTCATCA	3505
QY	539	AGTGGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGCCATCGGCACCG	598
Db	3506	AGTGGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGCCATCGGCACCG	3565
QY	599	TGCTGATCGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT	658
Db	3566	TGCTGGTGGGCGCCACCCCGTGAACATCATCGGCGCGCAACCTGCTGACCCAGATCGGCT	3625
QY	659	GCACCTTGAATTCGCCATCAGCCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCA	718
Db	3626	GCACCTTGAATTCGCCATCAGCCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCA	3685
QY	719	TGGACGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG	778
Db	3686	TGGACGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGTTGG	3745
QY	779	CCATCTCGAGGAGATGGAAGAGGAGGCAAGATCAACAGATCGGCCCGCGAGAACCCCT	838
Db	3746	AGATCTGCACCGAGATGGAAGAGGAGGCAAGATCAGCAAGATCGGCCCGAGAACCCCT	3805
QY	839	ACACACCCCGCTTTCGCCATCAAGAAAGAGAGACAGACCAAGTGGCGCAAGTGGTGG	898
Db	3806	ACACACCCCGCTTTCGCCATCAAGAAAGAGAGACAGACCAAGTGGCGCAAGTGGTGG	3865
QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC	958
Db	3866	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC	3925
QY	959	ACCCCGCGGCTTGAAGAAAGAGAGCGTGAACGCTGCTGGAGCTGGGCGACGCCTACT	1018
Db	3926	ACCCCGCGGCTTGAAGCAGAGAGAGCGTGAACGCTGCTGGAGCTGGGCGACGCCTACT	3985
QY	1019	TCAGCTGTCCTTGAGCAGAGACTTCCGCAAGTATACCCGCTTCACCATCCCCAGCATCA	1078
Db	5125	TCAGCTGTCCTTGAGCAGAGACTTCCGCAAGTATACCCGCTTCACCATCCCCAGCATCA	4045

QY 2159 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCAGCCGACAAAGAGGAGAGCG 2218
|||||
Db 5126 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCAGCCGACAAAGAGGAGAGCG 5185
QY 2219 AGCTGTGAACCAAGATCATCGACGAGCTGTATCAAGAGGAGAGAGGTACTAGCTGGG 2278
|||||
Db 5186 AGCTGTGAGCCAGATCATCGACGAGCTGTATCAAGAGGAGAGAGGTACTAGCTGGG 5245
QY 2279 TGCCCGCCACAGAGGAGGATCGCGGCAACGACGAGATCGACAGCTGGTGGACAGGGCA 2338
|||||
Db 5246 TGCCCGCCACAGAGGAGGATCGCGGCAACGACGAGGTGGACGCGCTGGTGGACGCGGCA 5305
QY 2339 TCCTGAAGGTGCTTCTTCTGGAGGAGGATCGATGGCGGCTGCTGATCTACCAAGTA 2393
|||||
Db 5306 TCCTGAAGGTGCTTCTTCTGGAGGAGGATCGACAAAGGCCAGGAGGACACAGAA 5360

RESULT 8
AX427927
LOCUS AX427927 12411 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 165 from Patent WO0232943.
ACCESSION AX427927
VERSION AX427927.1 GI:21538014
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.
1

REFERENCE
AUTHORS Huang, Y. and Nabel, G.-J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 165 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source location/Qualifiers
1. .12411
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2938 a 3733 c 3476 g 2264 t
ORIGIN

Query Match 82.1%; Score 2027; DB 6; Length 12411;
Best Local Similarity 91.6%; Pred. No. 2.1e-204;
Matches 2194; Conservative 0; Mismatches 185; Indels 16; Gaps 4;

QY 14 TGCCGAGGCCATGAGCCAGG---CCACCAGGCGCCACATCTCTGATCGAGGGCAGCAACT 70
|||||
Db 2967 TGCCGAGGCCATGAGCCAGGCTGACCAACAGCGGCCACCATCATGATCGAGCGGCAACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAGGAGGGCCACATCGGCC 130
|||||
Db 3027 TCCTCACACGCGCAAGATCGTGAAGTGTCTCAACTGCGGCAGGAGGGCCACACCGGCC 3086
QY 131 GCAACTCGCGGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCAAGAGGGCCACACAGA 190
|||||
Db 3087 GCAACTCGCGGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCAAGAGGGCCACACAGA 3146
QY 191 TGAGGACTGACCGAGCGGCCAGCCAACTTCTTCGCGAGGAGCTGGCTTCCCCCAGG 250
|||||
Db 3147 TGAAGGACTGACCGAGCGAGCAAGGCTAA-TTTTATTAGGAGATCTGGGCTTCCCAAG 3205
QY 251 GCAAGGCGCGGAGTTCCTCCAGCGCAGCAGAACCGCGCAACAGCCACAGCGCGCGAGC 310
|||||
Db 3206 GGAAGGCGAGGAATTTCTTCAGAGCAGACACAGAGCCACAGCCACACAGAGAGAGC 3255
QY 311 TGCAAGGTGCGCGG-----CGACAACCCCGCAGCGCGCGCGCGCGAGCGCCAGGGCA 364
|||||
Db 3266 TTCAGGTTTGGGGAAGAGACAACAACCTCTCTCAGAAAGCAGGAGCGGATAGACAAGAA 3325
QY 365 -----CCCTGAACCTTCCCGAGATCAACCTGTGGGAGCGGCCCTGGTGAAGCATCAAG 418
|||||
Db 3326 CTGATCTCTTAGCTTCTCCTCAGATCACTCTTTGGCAGCGACCCCTGCTCACAATAAGA 3385

QY 419 TGGCGGCCAGATCAAGGAGGGCCCTGCTGGACACCGCGCGACACACCGTGTCTGGAG 478
|||||
Db 3386 TAGGGGCGAGCTGAAGAGGCCCTTCTAGACACCGCGCGAGACACCGTGTCTGGAG 3445
QY 479 AGATGAGCCTTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGGCATCGGGGCTTCATCA 538
|||||
Db 3446 AGATGAACCTTGCCCGGCGCTTGAAGCCCAAGATGATCGGGGGCATCGGGGCTTCATCA 3505
QY 539 AGGTGCCCCAGTAGACAGATCCTGATCGAGATCTTGGGCAAGAAGGCCATCGGCACCG 598
|||||
Db 3506 AGGTGGGCCAGTAGACAGATCCTGATCGAGATCTTGGGCGCAAGGCCATCGGCACCG 3565
QY 599 TGCTGATCGGCCCCACCCCGCTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGCT 658
|||||
Db 3566 TGCTGGTGGGCCCCACCCCGCTGAACATCATCGGCCCGCAACATGCTGACCCAGATCGCT 3625
QY 659 GCACCTGAACCTTCCCATCAGCCCCATCGAGACCGTGGCCGTGCCGTGAAGTGAAGCCCGCA 718
|||||
Db 3626 GCACCTGAACCTTCCCATCAGCCCCATCGAGACCGTGGCCGTGAAGTGAAGCCCGCA 3685
QY 719 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAAGATCAAGGCCCTGACCG 778
|||||
Db 3686 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAAGATCAAGGCCCTGCTGG 3745
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 838
|||||
Db 3746 AGATCTGACCGAGATGGAGAAGGAGGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 3805
QY 839 ACAACACCCCGCTTCCCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGTGG 898
|||||
Db 3806 ACAACACCCCGCTTCCCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGTGG 3865
QY 899 ACTTCCCGGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC 958
|||||
Db 3866 ACTTCCCGGAGCTGAACAAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC 3925
QY 959 ACCCGCGCGGCTGAAGAAAGAGAGCGGTGACCGTGTCTGAGGTGGCGCACCGCTACT 1018
|||||
Db 3926 ACCCGCGCGGCTGAAGCAAGAGAGCGGTGACCGTGTCTGAGGTGGCGCACCGCTACT 3985
QY 1019 TCAGCGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCA 1078
|||||
Db 3986 TCAGCGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCA 4045
QY 1079 ACAAGAGACCCCGGATCGGCTACCAAGTACAGTGTCTGCCAGGCTGGAGGGCA 1138
|||||
Db 4046 ACAAGAGACCCCGGATCGGCTACCAAGTACAGTGTCTGCCAGGCTGGAGGGCA 4105
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTTGGAGCCCTTCCGCGCGCGCAACC 1198
|||||
Db 4106 GCGCCAGCATCTTCCAGTGCAGCATGACCAAGATCCTTGGAGCCCTTCCGCGAGCAACC 4165
QY 1199 CCGAGATCGTATCTACAGTACATGGACACCTGTAGTGGGAGCAGCAGCTGGAGATCG 1258
|||||
Db 4166 CCGACATCGTATCTACAGTACATGGACGACCTGTAGTGGGAGCAGCAGCTGGAGATCG 4225
QY 1259 GCCAGCAGCGCCCAAGATCGAGGAGTGCACAGCAGCTGTGGCTGGGCTTCACCA 1318
|||||
Db 4226 GCCAGCAGCGCCCAAGATCGAGGAGTGCACAGCAGCTGTGGCTGGGCTTCACCA 4285
QY 1319 CCGCCGCAAGAACACCAAGAGGAGCGCCCTTCTTGTGGATGGGTACGAGCTGCACC 1378
|||||
Db 4286 CCGCCGCAAGAACACCAAGAGGAGCGCCCTTCTTGTGGATGGGTACGAGCTGCACC 4345
QY 1379 CCGAAGTGAACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1438
|||||
Db 4346 CCGAAGTGAACCGTGCAGCCCATCGTGTCTGCCGAGAGACAGCTGGACCGTGAACG 4405
QY 1439 ACATCCAGAACTGGTGGGCAAGCTGAACTGGGCCAGCAGATCTACCCCGCATCAAG 1498
|||||
Db 4406 ACATCCAGAACTGGTGGGCAAGCTGAACTGGGCCAGCAGATCTACCCCGCATCAAG 4465

QY 1499 TGGCCACAGCTGTGCAAGCTGCTGGGGGCGCCCAAGGCCCTGACCGACATCGTCCGCCCTGA 1558
Db 4466 TGGCCACAGCTGTGCAAGCTGCTGGGGGCGCCCAAGGCCCTGACCGAGGTGGTCCGCCCTGA 4525
QY 1559 CCAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCCTGGCGAGCCCGTGCACG 1618
Db 4526 CCAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCCTGGAAGGAGCCCGTGCACG 4585
QY 1619 GCGTGTACTACGACCCACAGCAAGCACTGTGTGGCCGAGATCCCAAGAGCAGGCGCCAGAC 1678
Db 4586 GCGTGTACTACGACCCACAGCAAGCACTGTGTGGCCGAGATCCCAAGAGCAGGCGCCAGG 4645
QY 1679 AGTGGACCTACAGATCTACGAGAGGCCCTTCAAGAACCTGGAAGAGCAGGCGCCAGAC 1738
Db 4646 AGTGGACCTACAGATCTACGAGAGGCCCTTCAAGAACCTGGAAGAGCAGGCGCCAGAC 4705
QY 1739 AGATGCGCACCGCCACACAGCAAGCACTGTGTGGCCGAGATCCCAAGAGCAGGCGCCAG 1798
Db 4706 GATGAAGGCGGCCACACAGCAAGCACTGTGTGGCCGAGATCCCAAGAGCAGGCGCCAG 4765
QY 1799 CCATGGAGAGCATCGTGTGTCGGGCAAGACCCCAAGTTTCCGCCCTGCCCATCCAGAAGG 1858
Db 4766 CCACGAGAGCATCGTGTGTCGGGCAAGACCCCAAGTTTCAAGCTGCCATCCAGAGG 4825
QY 1859 AGACCTGGGAGACCTGGTGGACCGACTACTGGGAGCGCACTGGATCCCGAGTGGAGT 1918
Db 4826 AGACCTGGGAGGCTGGTGGACCGACTACTGGGAGCGCACTGGATCCCGAGTGGAGT 4885
QY 1919 TCCTGAACACCCCCCTGGTGAAGCTGTGTGACGCTGGAGAGGAGGAGCCCATCATCG 1978
Db 4886 TCCTGAACACCCCCCTGGTGAAGCTGTGTGACGCTGGAGAGGAGGAGCCCATCATCG 4945
QY 1979 GCGCCGAGACCTTCTACGTGGAGCGCGCCCAAGCCGAGACCAAGATCGGCAAGGCCG 2038
Db 4946 GCGCCGAGACCTTCTACGTGGAGCGCGCCCAAGCCGAGACCAAGATCGGCAAGGCCG 5005
QY 2039 GCTACGTGACCGAGCGGCGCGCAAGATCGTGAAGCTGACCGAGACCAAGATCGGCAAGGCCG 2098
Db 5006 GCTACGTGACCGAGCGGCGCGCAAGATCGTGAAGCTGACCGAGACCAAGATCGGCAAGGCCG 5065
QY 2099 AGACCGAGTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGGAGCGAGGTTGAACATCG 2158
Db 5066 AGACCGAGTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGGAGCGAGGTTGAACATCG 5125
QY 2159 TGACCGACAGCGAGTACGCGCTGGGCATCATCCAGGCCCGAGCCGACAAAGAGCGAGAGCG 2218
Db 5126 TGACCGACAGCGAGTACGCGCTGGGCATCATCCAGGCCCGAGCGAGGTTGAACATCG 5185
QY 2219 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2278
Db 5186 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 5245
QY 2279 TGCCCGCCCAAGGCGCATCGGCGGCAAGCAGAGCAGATCGACAGCTGGTGAAGGCGCA 2338
Db 5246 TGCCCGCCCAAGGCGCATCGGCGGCAAGCAGAGCAGTGTGAGCGGCTGTGAGGCGCGCA 5305
QY 2339 TCGGCAAGTGTGTCTTCTGGAGCGGATCGATGGGGGCACTCGTGTATCTACCACTA 2393
Db 5306 TCGGCAAGTGTGTCTTCTGGAGCGGATCGACAAAGGCCCGAGGAGGAGCAGAGAA 5360

RESULT 9
AX427926
LOCUS AX427926 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 164 from Patent WO0232943.
ACCESSION AX427926
VERSION AX427926.1 GI:21538013
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE artificial sequences.
1
AUTHORS Huang, Y. and Nabel, G. J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 164 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
1. 9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2672 c 2523 g 1767 t
ORIGIN
Query Match 82.0%; Score 2025.4; DB 6; Length 9194;
Best Local Similarity 91.6%; Pred. No. 3.4e-204;
Matches 2193; Conservative 0; Mismatches 186; Indels 16; Gaps 4;
QY 14 TGGCCGAGGCCATGAGCCAGG---CCACAGCGCCCAACATCTGATGTCAGCGGAGCAACT 70
Db 2967 TGGCCGAGGCCATGAGCCAGGTCACCAACAGCGCCACATCATGATGTCAGCGGAGCAACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCGCCACATCGCCCC 130
Db 3027 TCCGCAACACCGCCAGATCGTGAAGTGTCTCAACTGCGCAAGGAGGCGCCACCGCCCC 3086
QY 131 GCAACTGCCGCGCCCCCGCCCAAGAGGGCTGCTGGAAGTGCAGGAGGAGGCGCCACACAGA 190
Db 3087 GCAACTGCCGCGCCCCCGCCCAAGAGGGCTGCTGGAAGTGCAGGAGGAGGCGCCACACAGA 3146
QY 191 TGAAGACTGTCACCGAGCGCCAGCCAACTTCTTCGCGAGGAGCTGGCCCTTCCCCCAGG 250
Db 3147 TGAAGACTGTCACCGAGCGACAGGCTAA-TTTTATTAGGAGAGATCTGGCCCTTCCCCAAG 3205
QY 251 GCAAGCGCCGCGAGTTTCCCCAGCGAGCAGAACCCGCGCAACAGCCCGCCACAGCGCGGAGC 310
Db 3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACAGCCACAGCCCGCCACCAAGAGAGC 3265
QY 311 TGCAGGTGGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGAGCGCGCAGAGGCA 364
Db 3266 TTCAGGTTGGGGAAGAGACAACACTCCCTCTCAGAAGCAGGAGCGGATAGACAAGGAA 3325
QY 365 -----CCCTGAACTTCCCCAGATCACCTGTGGAGCGCCCGCTGGTAGCATCAAG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGAGCAGCGCCCTCGTGCACAAATAAGA 3385
QY 419 TGGCGCGCCAGATCAAGAGGCGCCCTGCTGGACACCGCGCGCGCAGCACACCGTCTGGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCCCTTCTAGACACCGCGCGCGCAGCACCGTCTGGAGG 3445
QY 479 AGATGAGCTGCCCGGCAAGTGGAGCCCAAGATGATCGGCGGCATCGGGCGGCTTCATCA 538
Db 3446 AGATGAACCTGGCCGCGCGCTGGAAGCCCAAGATGATCGGCGGCATCGGGCGGCTTCATCA 3505
QY 539 AGGTGCGCAGTACGACCGAGATCTGATCGAGATCTTCGCGCAAGAGGCCATCGGCACCG 598
Db 3506 AGGTGCGCAGTACGACCGAGATCTGATCGAGATCTTCGCGCAAGAGGCCATCGGCACCG 3565
QY 599 TGTGTATCGGCCCGCCACCCCGTGAACATCATCGGCCCGCAACATCTGACCGACTGGGCT 658
Db 3566 TGTGTGTGGGCCCGCCACCCCGTGAACATCATCGGCCCGCAACATCTGACCGACTGGGCT 3625
QY 659 GCACCTGAACTTCCCATCAGCCCCCATCGAGACCTGCCCCGTGAAGCTGGAAGCCCGGCA 718
Db 3626 GCACCTGAACTTCCCATCAGCCCCCATCGAGACCTGCCCCGTGAAGCTGGAAGCCCGGCA 3685
QY 719 TGGAGCGCCCGCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGAGCGCCCGCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 3745
QY 779 CCATCTGCGAGGAGATGGAAGAGGCGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 838
Db 3746 AGATCTGCGAGGAGATGGAAGAGGCGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 3805
QY 839 ACAACACCCCGCTGTTGCCCATCAAGAAAGACAGACCAAGTGGCGCAAGCTGGTGG 898

|||||
Db 3806 ACACACCCCGTTCGCCATCAAGAGAGAGACACCAAGTGGCGAAGCTGGTG 3865
QY 899 ACTTCGCGAGCTGACAAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCC 958
Db 3866 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAGAAGAGCGTGACCGTGTGGAGCTGGGGAGCGCTACT 1018
Db 3926 ACCCGCGCGCTGAAGAAGAAGAGCGTGACCGTGTGGAGCTGGGGAGCGCTACT 3985
QY 1019 TCAGCGTGCCTCGGAGGAGGACTTCGCAAGTACACCGCTTCACCATCCCCAGCATCA 1078
Db 3986 TCAGCGTGCCTCGGAGGAGGACTTCGCAAGTACACCGCTTCACCATCCCCAGCATCA 4045
QY 1079 ACACAGAGACCCCGGATCCGCTACCAAGTACAAAGTGTGCCCCAGGGGTGGAAGGCA 1138
Db 4046 ACAACAGAGACCCCGGATCCGCTACCAAGTACAAAGTGTGCCCCAGGGGTGGAAGGCA 4105
QY 1139 GCGCCAGCATCTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCGCCGCTATCTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CGGAGATCGTGATACCAAGTACATGAGACACCTGTACGTGGGAGGAGCTGGAGATCG 1258
Db 4166 CGGACATCGTGATACCAAGTACATGAGACACCTGTACGTGGGAGGAGCTGGAGATCG 4225
QY 1259 GCAGACCCCGGCAAGATCGAGAGGTGCGCAAGCAGCTGTGCGCTGGGGCTTACCA 1318
Db 4226 GCCAGCACCCGACAGATCGAGAGGTGCGCAAGCAGCTGTGCGCTGGGGCTTACCA 4285
QY 1319 CCGCCGACAAAGACACCAAGGAGCGCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 1378
Db 4286 CCGCCGACAAAGACACCAAGGAGCGCCCTTCTCTGTGGATGGGCTACGAGCTGCACC 4345
QY 1379 CGGACAAAGTGCAGCGTGCAGCGCTGCTGCTGCGGAGAGGAGCTGGACCGTGAACG 4405
Db 4405 ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGACCGCAGATCTACGCGCGCATCAAG 4465
QY 1499 TGGCGCAGCTGTCAAGCTGCTGCGGCGCCCAAGGCGCTGACGACATCGTGCCCTGA 1558
Db 4466 TGGCGCAGCTGTGCAAGCTGCTGCGGCGCCCAAGGCGCTGACGAGGTGTGCGCCCTGA 4525
QY 1559 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGCGGAGCCCGTGCAAG 1618
Db 4526 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAGGAGCCCGTGCAAG 4585
QY 1619 GCGTGTACTAGGACCCAGCAAGGACCTGTGGCGGAGATCCAGAACGAGGCGCACGACC 1678
Db 4586 GCGTGTACTAGGACCCAGCAAGGACCTGTGGCGGAGATCCAGAACGAGGCGCACGACC 4645
QY 1679 AGTGGACCTACCAAGTACCAAGGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 1738
Db 4646 AGTGGACCTACCAAGTACCAAGGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCA 4705
QY 1739 AGATGCGCACCGGCCACACCAAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG 1798
Db 4706 GCATGAAGGGGCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG 4765
QY 1799 CCATGGAGAGATCGTGATCTGGGCGAAGACCCCAAGTTCGCGCTCGCCATCCAGAAAG 1858
Db 4766 CCACCTGGAGAGATCGTGATCTGGGCGAAGACCCCAAGTTCGCGCTCGCCATCCAGAAAG 4825
QY 1859 AGACCTGGGAGACCTGTGTGACCGACTACTTGGCAGGCGACCTGTGATCCCGAGTGGGAGT 1918
Db 4826 AGACCTGGGAGGCTGTGTGACCGAGTACTTGGCAGGCGACCTGTGATCCCGAGTGGGAGT 4885
QY 1919 TCGTGAACACCCCGCTCGTGAAGCTGTGGTACCAAGCTGGGAGGAGGCCCATCATCG 1978
|||||

Db 4886 TCGTGAACACCCCGCTCGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCATCATCG 4945
QY 1979 GCGCGGAGACTTCTAGCTGGACGGCGCCGACCAACCGGAGAGACCAAGATGGCAAGGCCG 2038
Db 4946 GCGCGGAGACTTCTAGCTGGACGGCGCCGACCAACCGGAGAGACCAAGTGGCAAGGCCG 5005
QY 2039 GCTACGTGACCGACCGGGCGGCGAGAAAGTCTGTAGCCCTGACCGAGACCAACCAACGAGA 2098
Db 5006 GCTACGTGACCGACCGGGCGGCGAGAAAGTCTGTAGCCCTGACCGAGACCAACCAACGAGA 5065
QY 2099 AGACCGAGCTGCAGGCGCATCCAGCTGCGCCCTGACGAGACGCGGAGAGGTGAACATCG 2158
Db 5066 AGACCGAGCTGCAGGCGCATCCACCTGCGCCCTGACGAGACGCGGCTGGAGTGAACATCG 5125
QY 2159 TGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCGCCAGCCCGACAAAGAGCAGAGCG 2218
Db 5126 TGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCGCCAGCCCGACAAAGAGCAGAGCG 5185
QY 2219 AGCTGGTGAACACAGCATCATCGACGAGCTGTATCAAGAAGGAGAAAGTGTACCTGAGCTGGG 2278
Db 5186 AGCTGGTGAACACAGCATCATCGACGAGCTGTATCAAGAAGGAGAAAGTGTACCTGAGCTGGG 5245
QY 2279 TGCCCGCCCAAGAGGCGATCGCGGCAACGAGCAGATCGCAAGCTGTGTAGCAAGGGCA 2338
Db 5246 TGCCCGCCCAAGAGGCGATCGCGGCAACGAGCAGTGTGAGCGCTGTGTAGCGCGCGCA 5305
QY 2339 TCCGCAAGGTGCTGTCTCTGACGCGCATCGATGCGCGCATCGTGTATCATACAGTA 2393
Db 5306 TCCGCAAGGTGCTGTCTCTGACGCGCATCGCAAGGCGCCAGAGGAGCAGAGAA 5360

RESULT 10
AX427938
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 176 from Patent WO0232943.
ACCESSION AX427938
VERSION AX427938.1 GI:21538025
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 176 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source
1..9785
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2360 a 2843 c 2731 g 1851 t
ORIGIN

Query Match 81.8%; Score 2020.4; DB 6; Length 9785;
Best Local Similarity 91.6%; Pred. No. 1.1e-203;
Matches 2190; Conservative 0; Mismatches 181; Indels 21; Gaps 4;

QY 14 TGGCCGAGGCGCATGAGCCAGGCCACCGCCACATCTCTGATGCGCGCAGCAGCAACTTCA 73
Db 2967 TGGCCGAGGCGCATGAGCCAGGTGCGAGCACACCAACATCATGATGCGCGCGCAACTTCA 3026
QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGCGCACATCGGCCCGCA 133
Db 3027 AGGCGCCCAAGCG---CATCAAGTGTCTTCAACTGCGGCAAGGAGGCGCACCTTGGCCCGCA 3083
QY 134 ACTCCCGCGCCCGCCGAAAGAGGCTGTGGAAGTGTGCGCAAGGAGGCGCACACAGATGA 193
Db 3084 ACTCCCGCGCCCTCGGCAAGAGGCTGTGGAAGTGTGCGCAAGGAGGCGCACACAGATGA 3143
QY 194 AGGACTGACCGAGCGCGCAGGCCCAACTTCTTCCGCGAGGAGCTTGGCTTCCCCCGAGGCA 253
|||||

Db 3144 AGGACTGCACCGAGGACAGGCTAA-----TAGGGAAGATCTGGGCTTCCCCACAAGGGA 3197
QY 254 AGCGCCGCGAGTTCCCCAGCAGCAGAACCGCGCCCAACAGCCGCCACAGCGCGAGCTGC 313
Db 3198 AGCCAGGGAATTTTTCAGAGCAGCAGCAGACAGCAGCCGCCACCAAGAGAGCTTC 3257
QY 314 AGGTGCGCG- ----CGACACCCCGCAGCAGGCGCGCGCGCAGCGCCAGGGA--- 364
Db 3258 AGTTTGGGGAAGAGACAACAATCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAAGT 3317
QY 365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTCGTGGAGCATCAAGGTG 421
Db 3318 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAAGATAG 3377
QY 422 GCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCCGACACACCGCTGCTGGAGGAGA 481
Db 3378 GGGCCAGCTGAAGGAGGCCCTTCTAGACACCGCGCGCGAGACACCGTCTGTGAGGAGA 3437
QY 482 TGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGG 541
Db 3438 TGAACCTGCGCGCGCGCTGGAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGG 3497
QY 542 TGCGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCAACGTCG 601
Db 3498 TGGGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCAACGTCG 3557
QY 602 TGATCGGCGCCACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGTGCA 661
Db 3558 TGTGCGCGCCACCCCGTGAACATCATCGGCGCAACCTGCTGACCCAGATCGGCTGCA 3617
QY 562 CCCTGAACTTCCCGCATAGCCCATCGACCGTGCCTGCGTGAAGCTGAAGCCGCGGATGG 721
Db 3618 CCGTGAACTTCCCGCATAGCCCATCGACCGTGCCTGCGTGAAGCTGAAGCCGCGGATGG 3677
QY 722 ACGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACGCCCA 781
Db 3678 ACGGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGCTGGAGA 3737
QY 782 TCTGCGAGGAGATGGAGAAGGAGGGAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 841
Db 3738 TCTGCAACGAGATGGAGAAGGAGGGAAGATCAGCAAGATCGGCCCGCGAGAACCCCTACA 3797
QY 842 ACACCCCGCTGTCGCATCAAGAGAAGCAGACACCAAGTGGCGCAAGCTGGTGACAT 901
Db 3798 ACACCCCGCTGTCGCATCAAGAGAAGGAGCAGCACCAGTGGCGCAAGCTGGTGACAT 3857
QY 902 TCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGCAAGCTGGGCATCCCGCAC 961
Db 3858 TCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGCAAGCTGGGCATCCCGCAC 3917
QY 962 CCGCCGGCTGAAGAGAAGAGAGCGTGACCGTGTGGACGTGGCGACGCGCTACTTCA 1021
Db 3918 CCGCCGGCTGAAGCAGAGAAGAGCGTGACCGTGTGGACGTGGCGACGCGCTACTTCA 3977
QY 1022 GGTGCGCCCTGGACGAGGACTTCGCAAGTACACCGCTTCACCATCCCGCAGCATCAACA 1081
Db 3978 GGTGCGCCCTGGACAGGACTTCGCAAGTACACCGCTTCACCATCCCGCAGCATCAACA 4037
QY 1082 ACGAGACCCCGCGCATCCGCTACCAAGTACAAGGTGTGTCGCCAGGCGTGGAAAGGCGAGC 1141
Db 4038 ACGAGACCCCGCGCATCCGCTACCAAGTACAAGGTGTGTCGCCAGGCGTGGAAAGGCGAGC 4097
QY 1142 CCAGCATCTTCAGACAGCATGACCAAGATCCTGGAGCCCTTCGCGCGCCCGCAACCGCG 1201
Db 4098 CCGCCATCTTCAGTTCAGCATGACCATGACCAAGATCCTGGAGCCCTTCGCGCGCCCGCAACCGCG 4157
QY 1202 AGATCGTGATCTACAGTACATGGAGCGCTGTAGTGGGCGAGCGACCTGGAGATCGGCC 1261
Db 4158 ACATCGTGATCTACAGTACATGGAGCGCTGTAGTGGGCGAGCGACCTGGAGATCGGCC 4217
QY 1262 AGCAGCGCGCCAGATTCGAGGAGCTGCGCAAGCACCTGCTGGCGTGGGCTTCACCAACCC 1321
Db 4218 AGCAGCGCACCAAGATTCGAGGAGCTGCGCGACGACCTGCTGCGCTGGGCTTCACCAACCC 4277

QY 1322 CCGACAAGAGACCAAGAGAGCCCGCCCTTCTGTGGATGGCTACGAGCTGCACCCCG 1381
Db 4278 CCGACAAGAGACCAAGAGAGCCCGCCCTTCTGTGGATGGCTACGAGCTGCACCCCG 4337
QY 1382 ACAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAAGACA 1441
Db 4338 ACAAGTGGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAAGACA 4397
QY 1442 TCCAGAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCGGCATCAAGGTG 1501
Db 4398 TCCAGAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCGGCATCAAGGTG 4457
QY 1502 GCAGCTGTGCAAGCTGCTGCGCGGCCCAAGGCCCTGACCGACATCTGTCCTCCCTGACCG 1561
Db 4458 GCCAGCTGTGCAAGCTGCTGCGCGGCCCAAGGCCCTGACCGAGGTGGTCCCTGACCG 4517
QY 1562 AGAGCGCGAGCTGGAGCTGGCGCGAGAACCGCGAGATCTCTGCGCGAGCCGCTGACGGCG 1621
Db 4518 AGAGCGCGAGCTGGAGCTGGCGCGAGAACCGCGAGATCTCTGAAAGGAGCCGCTGACGGCG 4577
QY 1622 TGTACTACGACCCACGACAGGACCTGCTGGCCGAGATCCAGAGAGCGGCGACGACAGT 1681
Db 4578 TGTACTACGACCCACGACAGGACCTGATCGCCGAGATCCAGAGAGCGGCGACGAGT 4637
QY 1682 GGACCTTACCAGATCTTACCAGGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCCAAGA 1741
Db 4638 GGACCTTACCAGATCTTACCAGGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCCA 4697
QY 1742 TGGCGACCGCCACACACAGAGCTGAAGAGCTGACCGAGGCGCTGCGAGAAGATCGCCA 1801
Db 4698 TGAAGGCGCGCCACACACAGAGCTGAAGAGCTGACCGAGGCGCTGCGAGAAGATCGCCA 4757
QY 1802 TGGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCGCCCTGCCATCCAGAAAGAGA 1861
Db 4758 CCGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCAGCTGCGCCATCCAGAAAGAGA 4817
QY 1862 CTTGGAGAGCTTGGTGACCGACTTACTGGCAGGCCACCTTGGATCCCGCAGTGGGAGTTCG 1921
Db 4818 CTTGGAGAGCTTGGTGACCGAGTACTTGGCAGGCCACCTTGGATCCCGCAGTGGGAGTTCG 4877
QY 1922 TGAACACCCCGCCCTTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCATCATCGCG 1981
Db 4878 TGAACACCCCGCCCTTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCATCATCGCG 4937
QY 1982 CCGAGACCTTCTACGTGGACGGCGCCCAACCGCAGACCAAGATCGGCAAGCGCGCT 2041
Db 4938 CCGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGACCAAGCTGGGCAAGCGCGCT 4997
QY 2042 AGTGACCAACCGGGCGCGCAGAGATCTGAGCCTGACCGAGACCAACCAAGAGA 2101
Db 4998 ACGTGACCAACCGCGCGCGCAGAGGTGGTGGCCCTGACCGACACCAACCAAGAGA 5057
QY 2102 CCGAGCTGACGGCCATCCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCGTGA 2161
Db 5058 CCGAGCTGACGGCCATCCAGCTGGCCCTGCGAGGACAGCGGCTGGAGGTGAACATCGTGA 5117
QY 2162 CCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCCGCGCAGACAGAGGAGAGCGAGC 2221
Db 5118 CCGACAGCCAGTACGCCCTTGGGCATCATCCAGGCCCGCGCAGACAGAGGAGAGCGAGC 5177
QY 2222 TGGTGAACCAAGTATCTGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 2281
Db 5178 TGGTGAACCAAGTATCTGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 5237
QY 2282 CCGCCCAAGGCGCATCGGCGCAAGCAGCATCCAGAGCTGGTGAAGAGGGCATCC 2341
Db 5238 CCGCCCAAGGCGCATCGGCGCAAGCAGCATCCAGAGCTGGTGAAGAGGGCATCC 5297
QY 2342 GCAAGGTGCTTCTTGGAGCGCATCGATGGGGGCTCGTGTGATCTACCACTA 2393
Db 5298 GCAAGGTGCTTCTTGGAGCGCATCGATGGGGGCTCGTGTGATCTACCACTA 5349

RESULT 11
AX427921
LOCUS AX427921 9189 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 159 from Patent WO0232943.
ACCESSION AX427921
VERSION AX427921.1 GI:21538008
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 159 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1. 9189
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2671 c 2524 g 1762 t
ORIGIN
Query Match 81.8%; Score 2018.6; DB 6; Length 9189;
Best Local Similarity 91.4%; Pred. No. 1.7e-203;
Matches 2190; Conservative 0; Mismatches 184; Indels 21; Gaps 4;
Qy 14 TGGCCGAGGCATGAGCCAGG---CCACGAGCCCAACATCTTGATGAGCGCAGCAACT 70
Db 2967 TGGCCGAGGCATGAGCCAGGTCGAAGTCTTCAACTGGCGGAAGGAGGCGGCACATCGCCC 130
Qy 71 TCAAGGCCCCAAGCGCATCAAGTCTTCAACTGGCGGAAGGAGGCGGCACACAGA 190
Db 3027 TCCGCAACCAAGCGCATCAAGTCTTCAACTGGCGGAAGGAGGCGGCACACCGCC 3086
Qy 131 GCAACTGCGCGGCCCGCCGCAAGAGGCTGCTGGAGTGGCGGAAGGAGGCGGCACACAGA 190
Db 3087 GCAACTGCGCGGCCCGCCGCAAGAGGCTGCTGGAGTGGCGGAAGGAGGCGGCACACAGA 3146
Qy 191 TGAAGGACTGCACGAGCGCGCAGCGCAACTTCTTCCGCGAGGACTGCGCCCTTCCCGCAGG 250
Db 3147 TGAAGGACTGCACGAGCGCGCAGCGCAACTTCTTCCGCGAGGACTGCGCCCTTCCCGCAGG 3200
Qy 251 GCAAGGCGCGGAGTCTCCCGAGCGCAGACACCGCGCCACAGCGCCCGCAGCGCGAGC 310
Db 3201 GGAAGGCGCGGAGTCTTCTTCAAGCAGCAGACACCGCGCCACAGCGCCCGCAGCGAGC 3260
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGGAGCGCGCGGCA 364
Db 3261 TTCAGGTTTGGGAAGAGACACAACACTCCCTCTCAGRAGCAGAGCGCATAGACAAGAA 3320
Qy 365 -----CCCTGAACTTCCCGCAGATCAACCTGTGGCAGCGCCCGCTGTGAGCATCAAG 418
Db 3321 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGTCTCAATAAGA 3380
Qy 419 TGGCGCGCGCAGATCAAGGAGGCCCTGTGGACACCGCGCGCAGACACCTGTGTGAGG 478
Db 3381 TAGGGGCGCAGCTGAAGGAGGCCCTTTAGACACCGCGCGCAGACACCTGTGTGAGG 3440
Qy 479 AGATGAGCTTCCCGCGGAAGTGGAGCCCAAGATGATCGGGGGCATCGGGCGCTTCATCA 538
Db 3441 AGATGAACTTCCCGCGGCGGTGGGAAGCCCAAGATGATCGGGCGCATCGGGCGCTTCATCA 3500
Qy 539 AGGTGCGCCAGTACGACACAGATCTCTGATCGAGATCTGCGCAAGAAGGCCATCGGCACCG 598
Db 3501 AGGTGGGCGAGTACGACACAGATCTCTGATCGAGATCTGCGGCCACAAAGGCCATCGGCACCG 3560
Qy 599 TGCTGATGCGGCCCGCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3561 TGCTGGTGGGCGCCCGCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGATCGGCT 3620
Qy 659 GCACCTGAACCTTCCCATCAGCCCATCAGACCGTGCCTGAAGCTGAAGCCCGCA 718
Db 3621 GCACCTGAACCTTCCCATCAGCCCATCAGACCGTGCCTGAAGCTGAAGCCCGCA 3680
Qy 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGAACGAGGAGGATCAAGCCCTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGAACGAGGAGGATCAAGCCCTGCTGTG 3740
Qy 779 CCATCTCGAGGAGATGGAAGGAGGAGGAGGATCAACCAAGATCGCCCGCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAAGGAGGAGGAGGATCAACCAAGATCGCCCGCGAGAACCCCT 3800
Qy 839 ACAACACCCCGCTGTTCGCCCATCAAGAAGAGGACAGCACCACCAAGTGGCGCAAGCTGGTG 898
Db 3801 ACAACACCCCGCTGTTCGCCCATCAAGAAGAGGACAGCACCACCAAGTGGCGCAAGCTGGTG 3860
Qy 899 ACTTCGCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCAATCCCCC 958
Db 3861 ACTTCGCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCAATCCCCC 3920
Qy 959 ACCCGCGCGCTGAAGAAGAGAGAGAGCGTGCACCGTCTGGACGTGGCGACGCGCTACT 1018
Db 3921 ACCCGCGCGCTGAAGAAGAGAGAGAGCGTGCACCGTCTGGACGTGGCGACGCGCTACT 3980
Qy 1019 TCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGACCATCA 1078
Db 3981 TCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGACCATCA 4040
Qy 1079 ACAACGAGACCCCGGCTTCCGCTACCAAGTACAAGTCTGCCCGCAGGCTTGGAGGCA 1138
Db 4041 ACAACGAGACCCCGGCTTCCGCTACCAAGTACAAGTCTGCCCGCAGGCTTGGAGGCA 4100
Qy 1139 GCGCCGAGCATCTCCAGAGAGCATGACCAAGATCTTGAGGCGCTTCCGCGCGCCGCAAC 1198
Db 4101 GCGCGCGCATCTCCAGTGCAGCATGACCAAGATCTTGAGGCGCTTCCGCGCGCCGCAAC 4160
Qy 1199 CCGAGATCGTGATCTACCAAGTACATGAGACGCTTACGAGGAGCTTGGAGGAGTGG 1258
Db 4161 CCGAGATCGTGATCTACCAAGTACATGAGACGCTTACGAGGAGCTTGGAGGAGTGG 4220
Qy 1259 GCGAGCAGCGCGCAAGATCGAGGAGTGGCAGACCTGCTGCTGGGCTTGGGCTTCAACA 1318
Db 4221 GCGAGCAGCGCGCAAGATCGAGGAGTGGCAGACCTGCTGCTGGGCTTGGGCTTCAACA 4280
Qy 1319 CCGCGCAGACCAAGCAGCAGAGGAGGCGCCCTTCTGATGGGCTTGGGCTTGGGCTTCAACA 1378
Db 4281 CCGCGCAGACCAAGCAGCAGAGGAGGCGCCCTTCTGATGGGCTTGGGCTTGGGCTTCAACA 4340
Qy 1379 CCGACAAGTGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGGAGCTGGACCGTGAAG 1438
Db 4341 CCGACAAGTGGACCGTGCAGCGCATCGAGCTGCCGAGAGGAGGAGCTGGACCGTGAAG 4400
Qy 1439 ACATCCAGAAGCTGGTGGGCAAGTGAACCTGGGCGCAGCAGATCTTACCCCGCATCAAG 1498
Db 4401 ACATCCAGAAGCTGGTGGGCAAGTGAACCTGGGCGCAGCAGATCTTACCCCGCATCAAG 4460
Qy 1499 TGGCGCAGCTGTGAAGCTGTGCGCGCGCAGGCGCTTACGACATCTGTCGCCCTTGA 1558
Db 4461 TGGCGCAGCTGTGAAGCTGTGCGCGCGCAGGCGCTTACGACATCTGTCGCCCTTGA 4520
Qy 1559 CCGAGGAGGCGAGCTGGAGCTGGCGGAGACCGCGAGATCTGCGCGAGGCGCGTGGCAGG 1618
Db 4521 CCGAGGAGGCGAGCTGGAGCTGGCGGAGACCGCGAGATCTGCGCGAGGCGCGTGGCAGG 4580
Qy 1619 GCGTGTACTACGACCCCGCAGCAGGACCTTGGTGGCGGAGATCCAGAGCAGGCGCCACGACC 1678
Db 4581 GCGTGTACTACGACCCCGCAGCAGGACCTTGTGCGCGGAGATCCAGAGCAGGCGCCACGACC 4640
Qy 1679 AGTGGACCTACAGATCTTACGAGGAGCGCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 1738
Db 4641 AGTGGACCTACAGATCTTACGAGGAGCGCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 4700
Qy 1739 AGATGCGCACCGCCCGCACACCAACGACGTGAAGCAGCTGACCGAGGCGCGTGGCAGAGATCG 1798

Db 4038 ACAACGAGACCCCGGCATCGCTACCAAGTACAACGTGCTCCCGCAGGGCTGGAAGGSCA 4097
Qy 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACC 1198
Db 4098 GCCCGGCATCTTCCAGTGACATGACCAAGATCTCTGGAGCCCTTCGCGAAGCAGAAC 4157
Qy 1199 CCGAGATCGTGATCTACAGTACATGACGACCTGTGTACGTGGGAGGAGACCTGGAGATCG 1258
Db 4158 CCGACATCGTGATCTACAGTACATGACGACCTGTGTACGTGGGAGGAGACCTGGAGATCG 4217
Qy 1259 GCCAGCAGCGCCCAAGATCGAGGAGTGGCGAAGCAGCTGCTGGCTGGGGCTTCACCA 1318
Db 4218 GCCAGCAGCGCCCAAGATCGAGGAGTGGCGAAGCAGCTGCTGGCTGGGGCTTCACCA 4277
Qy 1319 CCCCAGACAAGACACGAGAGGAGGCCCCCTTCTCTGTGTGATGGGCTACGAGCTGCACC 1378
Db 4278 CCCCAGACAAGACACGAGAGGAGGCCCCCTTCTCTGTGTGATGGGCTACGAGCTGCACC 4337
Qy 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGGAGAGCTGGACGTGAACG 1438
Db 4338 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGGAGAGCTGGACGTGAACG 4397
Qy 1439 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCGGCATCAAG 1498
Db 4398 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCGGCATCAAG 4457
Qy 1499 TGCGCCAGCTGTCAAGCTGTGCGCGCGCCCAAGGCCCTGACGACATCGTGCCCTGA 1558
Db 4458 TGCGCCAGCTGTCAAGCTGTGCGCGCGCCCAAGGCCCTGACGAGGTGTGCCCTGA 4517
Qy 1559 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCTCGCGAGCCCGTGCAAG 1618
Db 4518 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCTCGCGAGCCCGTGCAAG 4577
Qy 1619 GCFTGTACTAGACCCCGAGAGAGCTGTGCGCGAGATCTCCAGAGCAGGCGCCAGCACC 1678
Db 4578 GCFTGTACTAGACCCCGAGAGAGCTGTGCGCGAGATCTCCAGAGCAGGCGCCAGGCGCC 4637
Qy 1679 AGTGAGCTACAGATCTACAGGAGCGCTTCAAGACCTCAAGACCGGCAAGTACGCCA 1738
Db 4638 AGTGAGCTACAGATCTACAGGAGCGCTTCAAGACCTCAAGACCGGCAAGTACGCCA 4697
Qy 1739 AGATGCCACCGGCCACACACGACGTGAAGCAGCTGACCGAGGCGGTGAGAGATCG 1798
Db 4698 GCATGAAGGCGCGCCACACACGACGTGAAGCAGCTGACCGAGGCGGTGAGAGATCG 4757
Qy 1799 CCATGGAGAGCATGTGATCTGGGGCAAGCCCGCAGTTCGGCTCGCCATCCAGAGG 1858
Db 4758 CCACCGAGAGCATGTGATCTGGGGCAAGCCCGCAGTTCAGCTGCCATCCAGAGG 4817
Qy 1859 AGACTGGGAGACTGTGTGACCGACTTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1918
Db 4818 AGACTGGGAGGCTGTGTGACCGAGTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 4877
Qy 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGCTGGAGAAGAGGCCCATCATCG 1978
Db 4878 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGCTGGAGAAGAGGCCCATCATCG 4937
Qy 1979 GCGCGGAGCTTCTAGCTGGAGCGCGCCGACCCACCGAGAGACCAAGATCGGCAAGCGCG 2038
Db 4938 GCGCGGAGCTTCTAGCTGGAGCGCGCGCCGACCCACCGAGAGACCAAGTGGGCAAGCGCG 4997
Qy 2039 GCTACGTGACCGACCGCGCGCGCAGAGATCTGTGAGCCTGACCGAGACCAACACAGAGA 2098
Db 4998 GCTACGTGACCGACCGCGCGCGCAGAGAGTGTGCCCCCTGACCGAGACCAACACAGAGA 5057
Qy 2099 AGACCGAGCTGCAGGCGCATCCAGTGGCCCTGCGAGGACAGCGGAGGTGAACATCG 2158
Db 5058 AGACCGAGCTGCAGGCGCATCCAGTGGCCCTGCGAGGACAGCGGCGTGAACATCG 5117
Qy 2159 TGACCGACAGCTAGCGCCCTGGGCATCATCCAGGCGCCGAGCCCGAGAGCGAGAGCG 2218
Db 5118 TGACCGACAGCTAGCGCCCTGGGCATCATCCAGGCGCCGAGCCCGAGAGCGAGAGCG 5177

Qy 2219 AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTACCTGAGCTGG 2278
Db 5178 AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTACCTGAGCTGG 5237
Qy 2279 TGCCCCGCCACAAAGGCGATCGCGGCAACGAGCAGATCGACAAGCTGTGTAGCAAGGCA 2338
Db 5238 TGCCCCGCCACAAAGGCGATCGCGGCAACGAGCAGTGTGTAGCAAGGCGCGCA 5297
Qy 2339 TCCCGCAAGTCTGTCTCTGTGACGGCATCGATGGCGGATCGTGTATCTACCACTA 2393
Db 5298 TCCCGCAAGTCTGTCTCTGTGACGGCATCGACAAGGCCCGAGGAGGACACAGAA 5352

RESULT 13
AX427928
LOCUS AX427928 9170 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 166 from Patent WO0232943.
ACCESSION AX427928
VERSION AX427928.1 GI:21538015
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
1 artificial sequences.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 166 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
Source 1..9170
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2225 a 2669 c 2519 g 1757 t
ORIGIN

Query Match 81.7% Score 2017; DB 6; Length 9170;
Best Local Similarity 91.4%; Pred. No. 2.6e-203;
Matches 2189; Conservative 0; Mismatches 185; Indels 21; Gaps 4;

Qy 14 TGGCCGAGGCGCATGAGCCAGG---CCACGAGCGCAACATCTGTATGTCAGCGCAGCAACT 70
Db 2967 TGGCCGAGGCGCATGAGCCAGGTCACCAACACGCGCCACCATCATGTATGTCAGCGCAGCAACT 3026
Qy 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGCGGCACATCGGCC 130
Db 3027 TCCGCAACCGCAGCGCAAGATCTGTAAGTCTTCAACTGCGGCAAGAGGCGGCACACCGGCC 3086
Qy 131 GCAACTCGCGCGCCCGCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGAGGCGGCACACAGA 190
Db 3087 GCAACTCGCGCGCCCGCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGAGGCGGCACACAGA 3146
Qy 191 TGAAGGACTGCACCGAGCGCGCCCAACTTCTTCCGCGAGGACCTGCTGCCCTTCCCGCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGCGCGCTAA-----TAGGAAGATCTGCGCTTCCCGCAGG 3200
Qy 251 GCAAGGCGCGGAGTTCCTCCAGCAGCAGAGAACCAGCGCAACAGCCCGCCAGCGCGCAGC 310
Db 3201 GGAAGGCGCGGAGTTCCTCCAGCAGCAGAGAACCAGCGCAACAGCCCGCCAGCGCGCAGC 3260
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCCAGCGAGGCGCGCGCGCGCGCGCGCGCA 364
Db 3261 TGCAGGTGCGCGG-----CGACAACCCCGCCAGCGAGGCGCGCGCGCGCGCGCGCGCA 3320
Qy 365 -----CCCTGAACTTCCCGCAGATCACCTTGTGGCAGCGCCCGCTGCTGAGCATCAAG 418
Db 3321 CTGTATCTTATGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCTGTACATATAAGA 3380
Qy 419 TGGCGCGCGAGATCAAGAGGCGCCCTGTGTGACACCGCGCGCGAGACACCGCTGTGAGG 478
Db 3381 TAGGGCGCGAGTGAAGGAGGCCCTTCTAGACACCGCGCGCGAGACACCGCTGTGAGG 3440

QY	479	AGATGAGCCTGCCCGCGCAAGTGGAAAGCCCAAGATGATCGCGGCATCGGGGCTTTCATCA	538
Db	3441		
		AGATGAACCTGCCCGCGCTGGAAAGCCCAAGATGATCGGGGCATCGCGGCTTCATCA	
QY	539	AGTGGCCGAGTACGACACAGATCTCTGATCGAGATCTCGGCGAAGAAGGCATCGGACCG	598
Db	3501		
		AGTGGCCGAGTACGACACAGATCTCTGATCGAGATCTCGGCGCAAGAAGGCATCGGACCG	
QY	599	TGCTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATCTCTCACTGCTCACTGGCT	658
Db	3561		
		TGCTGTGGCCCCACCCCGTGAACATCATCGGCCGCAACCTCTCTGACCAATCGGCT	
QY	659	GCACCTGAACCTCCCCATAGCCCCCATCGAGACCGTGCCTGAAAGCTGAAAGCCCGGCA	718
Db	3621		
		GCACCTGAACCTCCCCATAGCCCCCATCGAGACCGTGCCTGAAAGCTGAAAGCCCGGCA	
QY	719	TGACAGGCCCCCAAGGTGAAGCAGTGGCCCTTACCGGAGGAGAGATCAAGGCCCTGACCG	778
Db	3681		
		TGACAGGCCCCCAAGGTGAAGCAGTGGCCCTTACCGGAGGAGAGATCAAGGCCCTTGGTGG	
QY	779	CCATCTCGGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGAAGACCCCT	838
Db	3741		
		AGATCTGCACCGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGAAGACCCCT	
QY	839	ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGACCAAGTGCAGCGCAAGCTGTGG	898
Db	3801		
		ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGACCAAGTGCAGCGCAAGCTGTGG	
QY	899	ACTTCCGCGAGCTGACAGACGACCCAGGACTTCTGGAGTGCAGCTGGGCATCCCC	958
Db	3861		
		ACTTCCGCGAGCTGACAGACGACCCAGGACTTCTGGAGTGCAGCTGGGCATCCCC	
QY	959	ACCCCGCCGCGCTGAGAAGAAAGAGCGTGACCGTCTGAGCGTGGGAGAGCGCTACT	1018
Db	3921		
		ACCCCGCCGCGCTGAGAAGAAAGAGCGTGACCGTCTGAGCGTGGGAGAGCGCTACT	
QY	1019	TCAGCGTGGCCCTGGACGAGGACTTCCGCAAGTACACGCGCTTCACCATCCGACGATCA	1078
Db	3981		
		TCAGCGTGGCCCTGGACGAGGACTTCCGCAAGTACACGCGCTTCACCATCCGACGATCA	
QY	1079	ACAACGAGACCCCGGCATCCGCTACCAGTACAAGTCTGCCCCAGGCGTGAAGGGCA	1138
Db	4041		
		ACAACGAGACCCCGGCATCCGCTACCAGTACAAGTCTGCCCCAGGCGTGAAGGGCA	
QY	1139	GCCCCAGCATCTTCAGACGACGATGACACAGATCTCTGGAGCCCTTCCGGCCCGCAACC	1198
Db	4101		
		GCCCCAGCATCTTCAGTGCACATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAAC	
QY	1199	CCGAGATCTGTATACCAAGTACATGAGACCTGTAGCTGGGACGACGACTGGAGATCG	1258
Db	4161		
		CCGAGATCTGTATACCAAGTACATGAGACCTGTAGCTGGGACGACGACTGGAGATCG	
QY	1259	GCCAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCATCTCTCGCTGGGGGCTTCACCA	1318
Db	4221		
		GCCAGCACCGCCCAAGATCGAGGAGCTGCGCAAGCATCTCTCGCTGGGGGCTTCACCA	
QY	1319	CCCCGACAAAGACACAGAGAGAGCCCCCTTCCTGTGATGGGTACGAGCTGCACC	1378
Db	4281		
		CCCCGACAAAGACACAGAGAGAGCCCCCTTCCTGTGATGGGTACGAGCTGCACC	
QY	1379	CCGACAAGTGGACCGTGCAGCCCATCTAGCTGCCGGAAGGAGAGCTGGACCGTGAACG	1438
Db	4341		
		CCGACAAGTGGACCGTGCAGCCCATCTAGCTGCCGGAAGGAGAGCTGGACCGTGAACG	
QY	1439	ACATCCAGAAGCTGTGGCAAGCTGAACTGGGCGACGAGATCTACCCGGGCTCAAGG	1498
Db	4401		
		ACATCCAGAAGCTGTGGCAAGCTGAACTGGGCGACGAGATCTACCCGGGCTCAAGG	
QY	1499	TGGCCAGCTGTGCAGCTGTCTCGCGGGGCGCAAGGCCCTGACCGACATGTGTGCCCTGA	1558
Db	4461		
		TGGCCAGCTGTGCAGCTGTCTCGCGGGCACAAGGCCCTGACCGAGGTGTGTGCCCTGA	

Qy	1559	CCGAGGAGCCGAGCTGGAGCTGGCCGAGAACCCGCGAGATCCTTGGCGGAGCCCGCTGCACG	1618
Db	4521	CCGAGGAGCCGAGCTGGAGCTGGCCGAGAACCCGCGAGATCCTTGAAGGAGCCCGCTGCACG	4580
Qy	1619	CGCTGTACTACGACCCACGACAGGACCTTGGTGGCCGAGATCCAGAACGACGGGCCACGACC	1678
Db	4581	GCCTGTACTACGACCCACGACAGGACCTTATCCGCGAGATCCAGAACGACGGGCCACGAGCC	4640
Qy	1679	AGTGGACCTACCAAGATTACCAAGGAGCCCTTCAAGAACCTGAAGACCGGAAGTATACGCCA	1738
Db	4641	ATGAGGACCTACCAAGATTACCAAGGAGCCCTTCAAGAACCTGAAGACCGGAAGTATACGCCA	4700
Qy	1739	AGATGGCACCCGCCACACAAACGAGCTGAAGACGCTGACCGAGGCGGTGCAGAAGATCG	1798
Db	4701	GCATGAAGGGCGCCACACAAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCG	4760
Qy	1799	CCATGGAGACATCTGTATCTGGGCGAAGACCCCAAGTTCCGGCTGCCCATCCAGAAGG	1858
Db	4761	CCACCGAGACATCTGTATCTGGGCGAAGACCCCAAGTTCAAGTTCGCCCATCCAGAAGG	4820
Qy	1859	AGACCTGGGAGACCTGGTGACCGACTACTGTGCAGGCCACCTTGGATCCCCGAGTGGGAGT	1918
Db	4821	AGACCTGGGAGGCTTGGTGACCGAGTACTGCGAGGCCACCTTGGATCCCCGAGTGGGAGT	4880
Qy	1919	TCGTGAACACCCCCCTCGGTGAAGCTGTGTATCCAGCTGGAGAGGAGGCCATCATATCG	1978
Db	4881	TCGTGAACACCCCCCTCGGTGAAGCTGTGTATCCAGCTGGAGAGGAGGCCATCATATCG	4940
Qy	1979	GCGCCGAGACCTTCTACGTGGAGCGGCGCCGCAACCCGCGAGACCAAGATCGGCAAGCGCG	2038
Db	4941	GCGCCGAGACCTTCTACGTGGAGCGGCGCCGCAACCCGCGAGACCAAGTGGGCAAGCGCG	5000
Qy	2039	GCTACGTGACCGACCGGGCGGCGAGAAGATCGTGAGCGCTTACCCGAGACCAACCAACGAGA	2098
Db	5001	GCTACGTGACCGACCGGGCGGCGAGAAGTGGTGCCCTTGNCCGACACCAACCAACGAGA	5060
Qy	2099	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGAGGTGAACATCG	2158
Db	5061	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGAGGTGAACATCG	5120
Qy	2159	TGACCGACGCGAGTAGTGGCCCTTGGGCATCATCCAGGCCCCAGCCGACAAAGACGAGAGCG	2218
Db	5121	TGACCGACGCGAGTAGTGGCCCTTGGGCATCATCCAGGCCCCAGCCGACAAAGACGAGAGCG	5180
Qy	2219	AGCTTGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTTGAGCTGGG	2278
Db	5181	AGCTTGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTTGAGCTGGG	5240
Qy	2279	TGCCCCGCCACAGGGGATCGGGGCAACGACGACATGCACAGCTGGTGAAGCAGGGCCA	2338
Db	5241	TGCCCCGCCACAGGGGATCGGGGCAACGACGACATGCACAGCTGGTGAAGCAGGGCCA	5300
Qy	2339	TCGCAAGGTGCTTTCCTTGAGCGGCATCGATGGCGGCATCGTGATCTACCACTA	2393
Db	5301	TCCCAAGGTGCTTTCCTTGAGCGGCATCGATGGCGGCATCGATGGCGGCATCGATGGCGGC	5355
RESULT	14		
LOCUS	AX427935	9782 bp	DNA
DEFINITION	AX427935	Sequence 173 from Patent W00232943.	linear
ACCESSION	AX427935		
VERSION	AX427935.1	GI:21538022	
			PAT 20-JUN-

RESULT 14	DEFINITION	ORGANISM	REFERENCE	
AX427935	ACCESSION	SOURCE	AUTHORS	
LOCUS	VERSION	KEYWORDS	TITLE	
			JOURNAL	

FEATURES		Location/Qualifiers			
source		1..9782			
/organism="synthetic construct"					
/db_xref="taxon:32630"					
/note="plasmid pVR1012x/s containing HIV genes"					
BASE COUNT	2339 a	2895 c	2719 g	1829 t	
ORIGIN					
Query Match 81.7%; Score 2017; DB 6; Length 9782;					
Best Local Similarity 91.4%; Pred. No. 2.5e+203;					
Matches 2189; Conservative 0; Mismatches 185; Indels 21; Gaps 4;					
QY	14	TGGCGGAGGCGCATGAGCCAGG---CCACGACGGCCAAACATCCTGATGAGCGCGCAACT 70			
Db	3579	TGGCGAGGCGCATGAGCGAGGTGACCAAGCGCCACCATCATGATGAGCGCGCAACT 3638			
QY	71	TCAAGGGCCCCAAGCGCATCATGAAGTGTCTCACTGGCGCAAGGAGGGCCACATCGCCC 130			
Db	3639	TCCGCAACGCGCAAGATCGTGAAGTGTCTCACTGGCGCAAGGAGGGCCACACCGCCC 3698			
QY	131	GCAACTGCGCGCCCCCGCAAGAAGGCGTCTGTGAAGTGGCGAAGGAGGGCCACACAGA 190			
Db	3699	GCAACTGCGCGCCCCCGCAAGAAGGCGTCTGTGAAGTGGCGAAGGAGGGCCACACAGA 3758			
QY	191	TGAAGGACTGCACGAGGCGCCAGGCCAACTTCTTCCGCGAGGACCTGCGCTTCCGCCAGG 250			
Db	3759	TGAAGGACTGCACGAGGCGCCAGGCCAACTTCTTCCGCGAGGACCTGCGCTTCCGCCAGG 3812			
QY	251	GCAAGGCCCCGAGTTCCTCCAGCGGACGAGACACCGCGCCACAGCCCGCCAGCGAGC 310			
Db	3813	GGAAGGCCAGGGAATTTCTTTCAGAGCAGACCAAGCCAGAGCCACCGCCACAGAGAGC 3872			
QY	311	TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGAGGSCA 364			
Db	3873	TTCAAGTTTGGGAGAGACAACTCTCTTCAGAGCAGGAGCGCATAGACAAGGA 3932			
QY	365	-----CCCTGAATTCCTCCAGATCACTCTGTGGAGCGCCCTCGTGAGCATCAAG 418			
Db	3933	CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGA 3992			
QY	419	TGGCGGCGCAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGACACCTGTGTGAGG 478			
Db	3993	TAGGGGCGCAGCTGAAGAGGCGCTTCTAGACACCGGGCGCGAGACACCTGTGTGAGG 4052			
QY	479	AGATGAGCCTGCCCGGAAGTGGAGGCCAAGATGATCGGCGCATCGGCGCTTCATCA 538			
Db	4053	AGATGAACCTTSCCGCGCGCTGGAGCCCAAGATGATCGGCGCATCGGCGCTTCATCA 4112			
QY	539	AGGTGCGCAGTACGACCATCTGTATCGAGATCTGCGGCAAGAAGGCCATCGGCACCG 598			
Db	4113	AGGTGGGCGCAGTACGACCATCTGTATCGAGATCTGCGGCAAGAAGGCCATCGGCACCG 4172			
QY	599	TGCTGATCGGCGCCACCCCGTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 658			
Db	4173	TGCTGTGGGCGCCACCCCGTGAACATCATCGCGCCCAACATGCTGACCCAGATCGGCT 4232			
QY	659	GCACCTCTGAATTCCTCCATCAGCCCATCGAGACCGTGGCGGCAAGCTGAAGCCCGGCA 718			
Db	4233	GCACCTCTGAATTCCTCCATCAGCCCATCGAGACCGTGGCGGCAAGCTGAAGCCCGGCA 4292			
QY	719	TGGAGCGCCCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAAGATCAAGGCCCTGACCG 778			
Db	4293	TGGAGCGCCCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAAGATCAAGGCCCTTGTGG 4352			
QY	779	CCATCTCGAGGAGATGAGAGGAGGCGAGATCACCAGATCGGCGCCCGCAGAACCCCT 838			
Db	4353	AGATCTGCACGAGATGAGAGGAGGCGAGATCAGCAAGATCGGCGCCCGCAGAACCCCT 4412			
QY	839	ACAAACACCCCGTGTTCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGG 898			
Db	4413	ACAAACACCCCGTGTTCGCCATCAAGAAGAGGAGACAGCACCAAGTGGCGCAAGCTGTGG 4472			
QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGCACTGGGCAATCCCC 958			
Db	4473	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGCACTGGGCAATCCCC 4533			
QY	959	ACCCGCGCGGCTGAAGAAGAGAGAGCGTGACCGTGGTGGACGTGGCGACGCGCTACT 1018			
Db	4533	ACCCGCGCGGCTGAAGAAGAGAGAGCGTGACCGTGGTGGACGTGGCGACGCGCTACT 4592			
QY	1019	TCAGGTGCCCTGGACGAGGACTTCCGCAAGTACACGCGCTTCCACCTCCCGACGATCA 1078			
Db	4593	TCAGGTGCCCTGGACAGGACTTCCGCAAGTACACGCGCTTCCACCTCCCGACGATCA 4652			
QY	1079	ACAAGAGACCCCGGCTACCGTACCAAGTACCAAGTGTGCTCCCGCAGGCTGGAGGCA 1138			
Db	4653	ACAAGAGACCCCGGCTACCGTACCAAGTACCAAGTGTGCTCCCGCAGGCTGGAGGCA 4712			
QY	1139	GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGAAC 1198			
Db	4713	GCCCCAGCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAAG 4772			
QY	1199	CGGATCGTGATCTACCAAGTACATGGAGACCTTGTGCGGCGAGCGACCTGGAGATCG 1258			
Db	4773	CGGATCGTGATCTACCAAGTACATGGAGACCTTGTGCGGCGAGCGACCTGGAGATCG 4832			
QY	1259	GCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGCTGGGCTTTCACCA 1318			
Db	4833	GCCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGCTGGGCTTTCACCA 4892			
QY	1319	CCCCGCAAGAACGACCAAGAGGAGCCCGCTTCTGATGGATGGCTTACGAGCTGAC 1378			
Db	4893	CCCCGCAAGAACGACCAAGAGGAGCCCGCTTCTGATGGATGGCTTACGAGCTGAC 4952			
QY	1379	CCGACAAGTGACCGCTGACGCGCATCGAGCTGCGCGAGAGGAGAGCTGGACCTGAAG 1438			
Db	4953	CCGACAAGTGACCGCTGACGCGCATCGAGCTGCGCGAGAGGAGAGCTGGACCTGAAG 5012			
QY	1439	ACATCCAGAAGCTGGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGGCATCAAG 1498			
Db	5013	ACATCCAGAAGCTGGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCGCGGCATCAAG 5072			
QY	1499	TGCGGCGAGCTGTGCAAGCTGTGCGCGGCGGCGAGAGCTGCGGAGAGCTGCGGCG 1558			
Db	5073	TGCGGCGAGCTGTGCAAGCTGTGCGCGGCGGCGAGAGCTGCGGAGAGCTGCGGCG 5132			
QY	1559	CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGAGCTGCGGAGAGCTGCGGCG 1618			
Db	5133	CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGAGCTGCGGAGAGCTGCGGCG 5192			
QY	1619	CGGTGTACTACGACCCCGCAGCAAGGACCTGTGTGGCGGAGATCCAGAGAGGCGCCAG 1678			
Db	5193	CGGTGTACTACGACCCCGCAGCAAGGACCTGTATCGCGAGATCCAGAGAGGCGCCAG 5252			
QY	1679	AGTGACCTACAGATCTACCGAGGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACG 1738			
Db	5253	AGTGACCTACAGATCTACCGAGGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACG 5312			
QY	1739	AGATGCGCACCGCCACACCAACGAGCTGAAGCAGCTGACGAGGCGCGTGAGAGATCG 1798			
Db	5313	GCATGAAGGGCGCCACACCAACGAGCTGAAGCAGCTGACCGAGGCGCGTGAGAGATCG 5372			
QY	1799	CCATGGAGAGCATGTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 1858			
Db	5373	CCACCGAGAGCATGTGATCTGGGCAAGACCCCAAGTTCGAGCTGCCCATCCAGAGG 5432			
QY	1859	AGACCTGGGAGACCTGTGGAGCCGACTACTGGAGGCGCACCTGGATCCCCAGTGGAGT 1918			
Db	5433	AGACCTGGGAGCCCTGTGGAGCCGAGTACTGGCAGGCGCACCTGGATCCCCAGTGGAGT 5492			
QY	1919	TCGTGAACACCCCGCTGTGAGCTGTGTTACAGCTGGAGAGGAGCGCCATCATCG 1978			
Db	5493	TCGTGAACACCCCGCTGTGTAAGCTGTGTTACAGCTGGAGAGGAGCGCCATCATCG 5552			
QY	1979	CGCGCGAGACCTTCTACGTGGAGCGCGCGCAACCGGAGACCAAGATCGCAAGGCGG 2038			

Db	5553	CGCGCGAGACCTTCTACGTGGACGGCGCGGCCAACCGCGGAGACCAAGCTGGGCAAGGCCG	5612
Qy	2039	GCTACGTGACCGACCGGGCGCGCAGAAAGTCTGTAGCCTGACCGAGACCAACCAACAGCA	2098
Db	5613	GCTACGTGACCGACCGGGCGCGCAGAAAGTGTGTCCCTGACCGACACCAACCAACAGCA	5672
Qy	2099	AGACCGAGCTGCAGGGCCATCCAGCTGGCCCTGSCAGGACAGCGCGGACGAGTGAACATCG	2158
Db	5673	AGACCGAGCTGCAGGGCCATCCACCTGGCCCTGTCAGGACAGCGGCTGGAGGTGAACATCG	5732
Qy	2159	TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCCACCCGACAGAGCGAGCGG	2218
Db	5733	TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCCACCCGACAGAGCGAGCGG	5792
Qy	2219	AGCTGGTGAACACAGATCATCGACGAGCTGTATCAAGAGGAGAAGGTGTACCTGAGCTGGG	2278
Db	5793	AGCTGGTGAACACAGATCATCGACGAGCTGTATCAAGAGGAGAAGGTGTACCTGGCTGGG	5852
Qy	2279	TGCCCGGCCACAAGGGCATCGCGGCAACGACGAGATCGACAAGCTGGTGAGCAAGGGCA	2338
Db	5853	TGCCCGGCCACAAGGGCATCGCGGCAACGACGAGGTGGACGGCTGGTGAGCGCGGCA	5912
Qy	2339	TCCGCAAGGTGCTTCTCTGACGCGCATCGATGCGGCATCGTGATCTACCAGTA	2393
Db	5913	TCCGCAAGGTGCTTCTCTGACGCGCATCGACAAGGCGCCAGGAGGACACAGAA	5967
RESULT 15			
AX427934	AX427934	9783 bp	linear
LOCUS	AX427934	9783 bp	DNA
DEFINITION	Sequence 172 from Patent WO0232943.		PAT 20-JUN-2002
ACCESSION	AX427934		
VERSION	AX427934.1	GI:21538021	
KEYWORDS	synthetic construct.		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Huang Y. and Nabel G.J.		
TITLE	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization		
JOURNAL	Patent: WO 0232943-A 172 25-APR-2002;		
FEATURES	GOVERNMENT OF THE UNITED STATES (US)		
source	Location/Qualifiers		
	1..9783		
	/organism="synthetic construct"		
	/db.xref="taxon:32630"		
	/note="plasmid pVR1012x/s containing HIV genes"		
BASE COUNT	2339 a 2897 c 2723 g 1824 t		
ORIGIN			
Query Match	81.7%	Score 2017;	DB 6; Length 9783;
Best Local Similarity	91.4%;	Pred. No. 2.5e-203;	
Matches 2189;	Conservative 0;	Mismatches 185;	Indels 21; Gaps 4;
Qy	14	TGGCGAGGCCATGAGCCAGG---CCACGACGGCCAAACATCTGTATGATGAGCGGACGCAACT	70
Db	2964	TGGCGAGGCCATGAGCCAGGTGACCAACACGGCCACCATCATGATGAGCGCGGCAACT	3023
Qy	71	TCAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGGCCACATCGGCC	130
Db	3024	TCCGCAACCGCGCAAGATCTGTAAGTCTTCAACTGCGGCAAGGAGGGCCACACCGGCC	3083
Qy	131	GCACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA	190
Db	3084	GCAACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA	3143
Qy	191	TGAAGGACTGCACGAGCGCGCAGGCCAACTTCTTCCGGGAGGAGCTGGCTTCCCGCAGG	250
Db	3144	TGAAGGACTGCACGAGCGCGCAGGCTAA-----TAGGGAAGATCTGGCTTCCACAG	3197
Qy	251	GCAAGGCCCGCGAGTTCCCGCAGCGAGACCGGCCAACAGGCCCAACAGCGCGGAGC	310

...

Search completed: February 10, 2003, 17:55:41
Job time : 6389.22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:51:03 ; Search time 492.196 Seconds
(without alignments)
11296.684 Million cell updates/sec

Title: US-09-610-313-30
Perfect score: 2469
Sequence: 1 gtcgacgcaccatggccga.....gggtgacaccgtggaattc 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2469	100.0	2469	24	ABL39959
2	2442.2	98.9	2463	24	ABL39960
3	2415.4	97.8	2457	24	ABL39961
4	2052	83.1	2312	21	AAA70481
5	2051.2	83.1	9166	24	ABK91616
6	2042.8	82.7	9788	24	ABK91622
7	2028.8	82.2	9169	24	ABK91617
8	2027	82.1	9194	24	ABK91611
9	2027	82.1	12411	24	ABK91613
					Synthetic construc
					Synthetic construc
					Synthetic construc
					HIV FS(-).ProtMod
					Modified HIV prote
					Modified HIV prote
					Modified HIV prote
					Modified HIV prote
					Modified HIV prote

10	2025.4	82.0	9194	24	ABK91612
11	2025.2	82.0	2306	21	AAA70479
12	2020.4	81.8	9785	24	ABK91624
13	2018.6	81.8	9189	24	ABK91607
14	2017	81.7	9167	24	ABK91619
15	2017	81.7	9170	24	ABK91614
16	2017	81.7	9782	24	ABK91621
17	2017	81.7	9783	24	ABK91620
18	2017	81.7	9792	24	ABK91618
19	1999	81.0	3015	24	ABL39983
20	1998.4	80.9	2300	21	AAA70480
21	1982.4	80.3	3009	24	ABL40024
22	1969.4	79.8	4319	21	AAA70414
23	1935.8	78.4	8908	21	AA252055
24	1907	77.2	2577	24	AA141592
25	1905.8	77.2	4343	24	AA167896
26	1901.8	77.0	4341	24	AA167897
27	1881.2	76.2	1978	24	ABL39991
28	1879.2	76.1	7897	24	ABK91609
29	1858	75.3	2305	21	AAA70477
30	1837	74.4	1965	24	ABL39989
31	1831.2	74.2	2299	21	AAA70478
32	1815.6	73.5	3012	21	AA252054
33	1807.6	73.2	9407	24	ABK91623
34	1659.4	67.2	4307	20	AA208740
35	1659.4	67.2	4307	21	AAA93972
36	1659.4	67.2	4307	22	AAC86876
37	1651.4	66.9	4307	22	AAH43696
38	1651.4	66.9	4327	21	AAA93984
39	1651.4	66.9	4353	21	AAA93983
40	1651.4	66.9	4642	21	AAA93982
41	1651.4	66.9	9772	22	AAH43699
42	1608	65.1	1680	24	ABL39997
43	1592.8	64.5	4338	22	AAH22806
44	1592.8	64.5	8366	22	AAH22810
45	1554.4	63.0	1668	24	ABL39999

ALIGNMENTS

RESULT 1

ABL39959

ID ABL39959 standard; DNA; 2469 BP.

XX ABL39959;

XX ABL39959;

DT 15-MAY-2002 (first entry)

DE Synthetic construct PR975(+) SEQ ID NO:30.

DE Human immunodeficiency virus type C; antigenic HIV type C protein;

KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; nef;

KW immunostimulant; gene therapy; gene; ds.

XX Human immunodeficiency virus type C.

OS Synthetic.

XX WO200204493-A2.

XX 17-JAN-2002.

XX 05-JUL-2001; 2001WO-US21241.

XX 05-JUL-2000; 2000US-0610313.

XX (CHIR) CHIRON CORP.

PA (UYST-) UNIV STELLENBOSCH.

XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;

PI WPI; 2002-154920/20.

XX

PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -

XX Claim 1; Fig 8; 233pp; English.

CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (1). (1) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.

XX Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 other;

Query Match 100.0%; Score 2469; DB 24; Length 2469;
Best Local Similarity 100.0%; Pred No. 8, 2e-301;
Matches 2469; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GTGAGCGCCACCATGCGCGAGGCCATGAGCCAGSCCAGCCAGCCCAACATCCTGATGCAG	60
Db	1	GTGAGCGCCACCATGCGCGAGGCCATGAGCCAGSCCAGCCAGCCCAACATCCTGATGCAG	60
QY	61	CGCAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGCGCAGGAGGGC	120
Db	61	CGCAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGCGCAGGAGGGC	120
QY	121	CACATGCGCGCAACTTCCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTCGCGCAAGGAG	180
Db	121	CACATGCGCGCAACTTCCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTCGCGCAAGGAG	180
QY	181	GGCCACAGATGAAGACTGCACGAGCGCCAGGCCAATCTTCCGCGAGGACCTTGCC	240
Db	181	GGCCACAGATGAAGACTGCACGAGCGCCAGGCCAATCTTCCGCGAGGACCTTGCC	240
QY	241	TTCCCCCAGGGCAAGCGCGAGTTCCTCCAGCGAGCAGAACCGCGCCACAGCCCCACC	300
Db	241	TTCCCCCAGGGCAAGCGCGAGTTCCTCCAGCGAGCAGAACCGCGCCACAGCCCCACC	300
QY	301	AGCCGCGAGCTGCAGTGGCGGCGACACCCCGCAGCGAGCGCGCGCGAGCGCCAG	360
Db	301	AGCCGCGAGCTGCAGTGGCGGCGACACCCCGCAGCGAGCGCGCGCGAGCGCCAG	360
QY	361	GGCAGCTGAACCTTCCCGCAGATCACCCTGTGGCAGCGGCCCTGTGTAGCATCAAGTG	420
Db	361	GGCAGCTGAACCTTCCCGCAGATCACCCTGTGGCAGCGGCCCTGTGTAGCATCAAGTG	420
QY	421	GGCGGCGAGATCAAGAGGCGCTTGGACACCGCGCGCGAGCAGACACCTGTGGAGGAG	480
Db	421	GGCGGCGAGATCAAGAGGCGCTTGGACACCGCGCGCGAGCAGACACCTGTGGAGGAG	480
QY	481	ATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGCGCTTCATCAAG	540
Db	481	ATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGATCGCGCGCTTCATCAAG	540
QY	541	GTGCGCGAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGCGACCGTG	600
Db	541	GTGCGCGAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGCGACCGTG	600
QY	601	CTGATCGGCGCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGC	660
Db	601	CTGATCGGCGCCACCCCGTGAACATCATCGCGCGCGCAACATGCTGACCCAGCTGGGCTGC	660
QY	661	ACCTGAACCTTCCCATCATCGCGCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCATG	720
Db	661	ACCTGAACCTTCCCATCATCGCGCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCATG	720
QY	721	GACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGGC	780

Db	721	GACGGCCCCAAGTGAAGCAGTGGCCCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCC	780
QY	781	ATCTGGAGGAGATGGAGAAGGAGGCAAGATCACCAGATGGCCCCCGAGAGACCCCTAC	840
Db	781	ATCTGGAGGAGATGGAGAAGGAGGCAAGATCACCAGATGGCCCCCGAGAGACCCCTAC	840
QY	841	AACACCCCGCTGTTCGGCATCAAGAAGAAGAGAGCAGCAGCAAGTGGCGCAAGCTGGTGC	900
Db	841	AACACCCCGCTGTTCGGCATCAAGAAGAAGAGAGCAGCAGCAGCAAGTGGCGCAAGCTGGTGC	900
QY	901	TTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCGATCCCCCAG	960
Db	901	TTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGCGATCCCCCAG	960
QY	961	CCCGCGGCTGGAAGAAGAAGAGAGGCGTACCGTGTGAGAGTGGCGGAGCGCTACTTC	1020
Db	961	CCCGCGGCTGGAAGAAGAAGAGAGGCGTACCGTGTGAGAGTGGCGGAGCGCGCTACTTC	1020
QY	1021	AGCGTGGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTACCATCCCCCAGCATCAAC	1080
Db	1021	AGCGTGGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTACCATCCCCCAGCATCAAC	1080
QY	1081	AACGAGACCCCGCGCATCCGCTACCAAGTACAGCTGTGCTGCCCGCAGGCGTGAAGGCG	1140
Db	1081	AACGAGACCCCGCGCATCCGCTACCAAGTACAGCTGTGCTGCCCGCAGGCGTGAAGGCG	1140
QY	1141	CCGAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGGCGCTTCCGCGCGCGCAACCC	1200
Db	1141	CCGAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGGCGCTTCCGCGCGCGCAACCC	1200
QY	1201	GAGATCGTGTATACCAAGTACATGGAGAGCTGTACCTGGGCGAGCAGCTGGAGATCGG	1260
Db	1201	GAGATCGTGTATACCAAGTACATGGAGAGCTGTACCTGGGCGAGCAGCTGGAGATCGG	1260
QY	1261	CAGCACCAGGCGCAAGTTCGAGGAGCTGCGCAGACCTGCTGCGCTGGGCGTTCACCAAC	1320
Db	1261	CAGCACCAGGCGCAAGTTCGAGGAGCTGCGCAGACCTGCTGCGCTGGGCGTTCACCAAC	1320
QY	1321	CCGCAAGAAGCAGCAGAGGAGCGCCCTTCTGCTGGATGGGCTACGAGCTGCACCCC	1380
Db	1321	CCGCAAGAAGCAGCAGAGGAGCGCCCTTCTGCTGGATGGGCTACGAGCTGCACCCC	1380
QY	1381	GACAAGTGAAGCTGAGCGCATCGAGCTGCGCGAGAGAGAGCTGAGCGCTGAACGAG	1440
Db	1381	GACAAGTGAAGCTGAGCGCATCGAGCTGCGCGAGAGAGAGCTGAGCGCTGAACGAG	1440
QY	1441	ATCCAGAGCTGTGGGCAAGCTGAACCTGGGCGAGCGAGATCTACCCCGCATCAAGGTG	1500
Db	1441	ATCCAGAGCTGTGGGCAAGCTGAACCTGGGCGAGCGAGATCTACCCCGCATCAAGGTG	1500
QY	1501	CGCGAGCTGTGAAGCTGTGCGCGGCGCAAGGCCCTGACCGACATCTGCCCCCTGACC	1560
Db	1501	CGCGAGCTGTGAAGCTGTGCGCGGCGCAAGGCCCTGACCGACATCTGCCCCCTGACC	1560
QY	1561	GAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCGCGCTGCACGG	1620
Db	1561	GAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCGCGCTGCACGG	1620
QY	1621	GTGTACTAGACCCCGCAGAGGAGCTGGTGGCGGAGATCCAGAGGAGGCGCCAGACG	1680
Db	1621	GTGTACTAGACCCCGCAGAGGAGCTGGTGGCGGAGATCCAGAGGAGGCGCCAGACG	1680
QY	1681	TGGACCTACCATCTTACAGAGGAGCTGAGAGCTGACCGAGGCGCTGCAAGAGATCGCC	1740
Db	1681	TGGACCTACCATCTTACAGAGGAGCTGAGAGCTGACCGAGGCGCTGCAAGAGATCGCC	1740
QY	1741	ATGCGCAGCGCCACACACGAGCTGAGAGCTGACCGAGGCGCTGCAAGAGATCGCC	1800
Db	1741	ATGCGCAGCGCCACACACGAGCTGAGAGCTGACCGAGGCGCTGCAAGAGATCGCC	1800
QY	1801	ATGAGGAGCATCTGTATCTGGGCGCAAGACCCCAAGTTCGCGCTGCGCATCTCAGAGAG	1860
Db	1801	ATGAGGAGCATCTGTATCTGGGCGCAAGACCCCAAGTTCGCGCTGCGCATCTCAGAGAG	1860

Db 1801 ATGAGAGCATCGTGATCGGGCAGAGCCCAAGATTCCGCTGCCCATCCAGAAGGAG 1860
Qy 1861 ACCTGGAGACCTGTGGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGAGTTC 1920
Db 1861 ACCTGGAGACCTGTGGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGAGTTC 1920
Qy 1921 GTGAACACCCCCCTGTGAAGCTGTGTACCAAGCTGGAGAGGAGCCATCATCGGC 1980
Db 1921 GTGAACACCCCCCTGTGAAGCTGTGTACCAAGCTGGAGAGGAGCCATCATCGGC 1980
Qy 1981 GCCGAGACCTTCTAGCTGGAGGCCGCCGCCCAACCCCGAGACAAGATCGGCAAGCCCGC 2040
Db 1981 GCCGAGACCTTCTAGCTGGAGGCCGCCGCCCAACCCCGAGACAAGATCGGCAAGCCCGC 2040
Qy 2041 TACGTGACCGACCGGGCGCGGAGAGATCGTGAGCCCTGACGAGACACCAACAGAG 2100
Db 2041 TACGTGACCGACCGGGCGCGGAGAGATCGTGAGCCCTGACGAGACCAACAGAG 2100
Qy 2101 ACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGAGGAGTGAACATCGTG 2160
Db 2101 ACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACGCGAGGAGTGAACATCGTG 2160
Qy 2161 ACCGAGACCGAGTACGGCTTGGCATCATCCAGGCCAGCCGACAAAGCGAGAGCCAG 2220
Db 2161 ACCGAGACCGAGTACGGCTTGGCATCATCCAGGCCAGCCGACAAAGCGAGAGCCAG 2220
Qy 2221 CTGTTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGGTTACTGAGCTGGGTG 2280
Db 2221 CTGTTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGGTTACTGAGCTGGGTG 2280
Qy 2281 CCGGCCCAACAGGGCATCGCGGCAACGAGCAGATCGACAGTGGTGAGCAAGGGCATC 2340
Db 2281 CCGGCCCAACAGGGCATCGCGGCAACGAGCAGATCGACAGTGGTGAGCAAGGGCATC 2340
Qy 2341 CGCAGGTGCTGTCCTGAGCGGCATCGATCGCGCATCGTGATCTACCATGATGAC 2400
Db 2341 CGCAGGTGCTGTCCTGAGCGGCATCGATCGCGCATCGTGATCTACCATGATGAC 2400
Qy 2401 GACCTGTACGTGGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2460
Db 2401 GACCTGTACGTGGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2460
Qy 2461 GGTGAATTC 2469
Db 2461 GGTGAATTC 2469

RESULT 2

ABL39960
ID ABL39960 standard; DNA; 2463 BP.
XX ABL39960;
XX
XX
XX
DT 15-MAY-2002 (first entry)
XX
XX
DE Synthetic construct PR975YM SEQ ID NO:31.
XX
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
XX
XX WO200204493-A2.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US21241.
XX
XX 05-JUL-2000; 2000US-0610313.
XX
XX (CHIR) CHIRON CORP.
PA

PA

(UYST-) UNIV STELLENBOSCH.

Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
WPI; 2002-154920/20.

New polynucleotides encoding antigenic HIV Type C polypeptides, useful
in applications including DNA immunization or generation of packaging
cell lines, particularly in gene therapy -

Claim 1; Fig 9; 233pp; English.

The present invention describes expression cassettes comprising a
polynucleotide sequence encoding a polypeptide comprising immunogenic
HIV type C polypeptides. The expression cassettes comprise any of the
HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
Nef (i). (i) have immunostimulant activity and can be used in gene
therapy. The HIV type C polynucleotides are useful in applications
including DNA immunisation, generation of packaging cell lines, and
production of HIV Type C proteins. The polynucleotides are particularly
useful in gene therapy and DNA immunisation applications. ABL39942 to
ABU40054 and ABB06204 to ABB06215 represent sequences used in the
exemplification of the present invention.

Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 other;

Query Match 98.9%; Score 2442.2; DB 24; Length 2463;

Best Local Similarity 99.6%; Pred. NO. 1.9e-297;

Matches 2460; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTCGACGCCACCATGGCGAGGCCATGAGCCAGGCCAGCCAGCCCAACTCTCTGATGCGAG 60

Db 1 GTCGACGCCACCATGGCGAGGCCATGAGCCAGGCCAGCCAGCCCAACTCTCTGATGCGAG 60

Qy 61 GCGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGGGCAAGGAGGCG 120

Db 61 GCGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGGGCAAGGAGGCG 120

Qy 121 CACATCGCCCGCAACTGCGGCGCCCCCGCCAGAGAGGGCTGCTGGAAGTGGCGGCAAGGAG 180

Db 121 CACATCGCCCGCAACTGCGGCGCCCCCGCCAGAGAGGGCTGCTGGAAGTGGCGGCAAGGAG 180

Qy 181 GCGCACCATGAGGACTGCAACCGAGCGCCAGGCGCAACTTCTCCGAGGAGCACTGGCC 240

Db 181 GCGCACCATGAGGACTGCAACCGAGCGCCAGGCGCAACTTCTCCGAGGAGCACTGGCC 240

Qy 241 TTCCGCCAGGCGAAGGCCCGCGAGTTCCCGAGCAGAGAACCGCGCCCAACAGCCCAACC 300

Db 241 TTCCGCCAGGCGAAGGCCCGCGAGTTCCCGAGCAGAGAACCGCGCCCAACAGCCCAACC 300

Qy 301 AGCGCGAGCTGCAGGTGCGGCGGCACAAACCCCGCAGGAGGCCGCGCGAGCGCCAG 360

Db 301 AGCGCGAGCTGCAGGTGCGGCGGCACAAACCCCGCAGGAGGCCGCGCGAGCGCCAG 360

Qy 361 GGCACCTTGAATCTCCCGCAGATCACCTGTGGCAGCGCCCGCTGCTGAGCATCAAGGTG 420

Db 361 GGCACCTTGAATCTCCCGCAGATCACCTGTGGCAGCGCCCGCTGCTGAGCATCAAGGTG 420

Qy 421 GCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGACACACCTGCTGGAGGAG 480

Db 421 GCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGACACACCTGCTGGAGGAG 480

Qy 481 ATGAGCCTGCCCGCAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCTATCAAG 540

Db 481 ATGAGCCTGCCCGCAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCTATCAAG 540

Qy 541 GTGGCCAGTACGACCATGATCCTGATCGAGATCTGCGGCAAGAGGCCCATCGGACCCGTG 600

Db 541 GTGGCCAGTACGACCATGATCCTGATCGAGATCTGCGGCAAGAGGCCCATCGGACCCGTG 600

Qy 601 CTGATCGGCCCGCCCGCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGC 660

Db 601 CTGATCGGCCCGCCCGCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGC 660

[illegible]

Db 1609 GTGTACTACGACCCCAAGACCTGGTGGCGAGATCCAGAAAGCAGGGCCACGACGAG 1668
 Qy 1681 TGGACCTACAGATCTACCAAGGAGCCCTTCAAGAACTGAAGACCGCAAGTACGCCAAG 1740
 Db 1669 TGGACCTACAGATCTACCAAGGAGCCCTTCAAGAACTGAAGACCGCAAGTACGCCAAG 1728
 Qy 1741 ATGGCACCACCCACACCAACGACGCTGAAGCAGCTGACCCAGGCGCTGCAGAAAGATCGCC 1800
 Db 1729 ATGGCACCACCCACACCAACGACGCTGAAGCAGCTGACCCAGGCGCTGCAGAAAGATCGCC 1788
 Qy 1801 ATGGAGAGCATCGTGTCTGGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAAGAG 1860
 Db 1789 ATGGAGAGCATCGTGTCTGGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAAGAG 1848
 Qy 1861 ACCTGGGAGACCTGGTGGACCGGCTACTGTGACGACGACCTGGATCCCGAGTGGGAGTTC 1920
 Db 1849 ACCTGGGAGACCTGGTGGACCGGCTACTGTGACGACGACCTGGATCCCGAGTGGGAGTTC 1908
 Qy 1921 GTGAACACCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGC 1980
 Db 1909 GTGAACACCCCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGC 1968
 Qy 1981 GCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2040
 Db 1969 GCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2028
 Qy 2041 TACGTGACCGACCGGCGCGGCGGAGAGATCGTGAGCCTGACCGAGACCAACCAAG 2100
 Db 2029 TACGTGACCGACCGGCGCGGCGGAGAGATCGTGAGCCTGACCGAGACCAACCAAG 2088
 Qy 2101 ACCGAGCTGACGAGCCATCCAGCTGGCCCTGACGAGACGCGGACGAGGTGAACATCGTG 2160
 Db 2089 ACCGAGCTGACGAGCCATCCAGCTGGCCCTGACGAGACGCGGACGAGGTGAACATCGTG 2148
 Qy 2161 ACCGAGACGAGTACGCGCTGGGCATCATCCAGCCCGACCCGACAGAGCGAGCGAG 2220
 Db 2149 ACCGAGACGAGTACGCGCTGGGCATCATCCAGCCCGACCCGACAGAGCGAGCGAG 2208
 Qy 2221 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTG 2280
 Db 2209 CTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTG 2268
 Qy 2281 CCCGCCACAAAGGAGATCGGCGGCAAGCAGAGATCGACAAAGCTGGTGAGCAAGGGCATC 2340
 Db 2269 CCCGCCACAAAGGAGATCGGCGGCAAGCAGAGATCGACAAAGCTGGTGAGCAAGGGCATC 2328
 Qy 2341 CGCAAGCTGCTTCTCTGGAGGCGATCGATGGGCGCATCGTGATCTACCAAGTACATGAC 2400
 Db 2329 CGCAAGCTGCTTCTCTGGAGGCGATCGATGGGCGCATCGTGATCTACCAAGTACATGAC 2388
 Qy 2401 GACCTGTACGTGGGAGCGGCGGCTTAGGATCGATTAAAGCTTCCCGGGGCTTAGCAC 2460
 Db 2389 GACCTGTACGTGGGAGCGGCGGCTTAGGATCGATTAAAGCTTCCCGGGGCTTAGCAC 2448
 Qy 2461 GGTGAATTC 2469
 Db 2449 GGTGAATTC 2457

RESULT 4

AAA70481
 ID AAA70481 standard; DNA; 2312 BP.

XX
 AC AAA70481;

DT 28-NOV-2000 (first entry)

XX HIV FS(-)_ProtMod_Rtopt(+) coding sequence.

DE HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.

KW HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.

XX Human immunodeficiency virus type 1.

OS Synthetic.

XX WO200039302-A2.
 PN 06-JUL-2000.
 PD 30-DEC-1999; 99WO-US31245.
 PF 31-DEC-1998; 98US-0114495.
 PR 01-DEC-1999; 99US-0168471.
 XX (CHIR) CHIRON CORP.
 PA Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
 PI Greer C, Selby M, Walker C;
 XX WPI; 2000-452400/39.
 DR
 XX Expression cassettes encoding the human immunodeficiency virus (HIV)
 PT Gag-containing polypeptide useful for vaccinating against HIV
 PT Infections and acquired immunodeficiency syndrome (AIDS) -
 XX Claim 7; Fig 75; 391pp; English.
 XX The present sequence is the coding sequence of a synthetic HIV
 CC Gag-reverse transcriptase expression cassette, FS(-)_ProtMod_Rtopt(+).
 CC The Gag protein of HIV is needed for the assembly of virus-like
 CC particles. In addition, the Gag protein is involved in many stages of the
 CC HIV life cycle, including assembly, virion maturation after particle
 CC release and early post-entry steps in viral replication. The expression
 CC cassette may be used for the recombinant expression of HIV
 CC Gag-polypeptides which may then be used to vaccinate against HIV
 CC infection and acquired immunodeficiency syndrome (AIDS).
 XX Sequence 2312 BP; 533 A; 750 C; 732 G; 297 T; 0 other;

Query Match 83.1%; Score 2052; DB 21; Length 2312;
 Best Local Similarity 93.6%; Pred. No. 1.le-248;
 Matches 2165; Conservative 0; Mismatches 135; Indels 12; Gaps 2;

Qy 170 GCGCAAGGAGGCGCACGATGAAGACTGCACCGGAGCGGCGCAACTTCTTCCGCG 229
 Db 1 GCGGCCCGAAGGACACCAATGAAGATTGCATGAGACAGGCTAATTTCTCCGCG 60
 Qy 230 AGGACCTGGGCTTCCCGCAGGGCAAGCGCGAGTTCCCGCAGGAGCAACCGCGCA 289
 Db 61 AGGACCTGGGCTTCTGCAGGGCAAGCGCGAGTTTCAGCAGGAGCAGACCGCGCA 120
 Qy 290 ACAGCCCCACCGCGAGCTGCAGTGCAGGCGCGGCGGCGGCGGCGGCGGCGG 343
 Db 121 ACAGCCCCACCGCGAGCTGCAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 Qy 344 CCGCGCGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 397
 Db 181 CCGCGCGCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Qy 398 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 457
 Db 241 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
 Qy 458 CCGAGCAGACCGTCTGGAGGAGATGAGCCCTGCGCGGCGGCGGCGGCGGCGGCGG 517
 Db 301 CCGAGCAGACCGTCTGGAGGAGATGAGCCCTGCGCGGCGGCGGCGGCGGCGGCGG 360
 Qy 518 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 577
 Db 361 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
 Qy 578 GCAGAGAGGCGCATCGGACCGCTGCTGATCGGCGGCGGCGGCGGCGGCGGCGGCGG 637
 Db 421 GCCAAGAGGCGCATCGGACCGCTGCTGATCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
 Qy 638 ACATGCTGACCGAGCTGGGCTGACCGCTGAACTTCCCGCAGCGGCGGCGGCGGCGG 697

Db 481 ACCTGCTGACCCAGATCGGCTGCACCCCTGAACCTTCCCATCAGCCCCATCGAGACGGTGC 540
QY 698 CGTGAAGCTGAAGCCCGCATGGACGGCCCAAGGTGAAGAGTGGCCCTTGACCCGAGG 757
Db 541 CCGTGAAGCTGAAGCCCGGATGGACGGCCCAAGGTGAAGAGTGGCCCTTGACCCGAGG 600
QY 758 AGAAGATCAAGGCCCTGACCCGCTCTGCGAGGAGATGGAGAGGAGGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTGCTGGAGATCTGCACCGAGATGGAGAGGAGGCAAGATCACCA 660
QY 818 AGATCGGCCCGAGAAACCCCTACACACCCCGGTGTTGCGCCATCAAGAAAGAGACAGCA 877
Db 661 AGATCGGCCCGAGAAACCCCTACACACCCCGGTGTTGCGCCATCAAGAAAGAGACAGCA 720
QY 878 CCAAGTGGCGAAGCTGTGTGACTTCCGCGAGCTCAACAAAGCGCACCCAGGACTTCTGGG 937
Db 721 CCAAGTGGCGAAGCTGTGTGACTTCCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGG 780
QY 938 AGGTGCAGCTGGGCATCCCCACCCCGCCGCTTGAAGAAAGAGAGCGCTGACCCGTGC 997
Db 781 AGGTGCAGCTGGGCATCCCCACCCCGCCGCTTGAAGAAAGAGAGCGCTGACCCGTGC 840
QY 998 TGGAGTGGGCGACCCCTACTTTCAGCTGCCCCCTGGAGAGAGCTTCGCGAAGTACACCG 1057
Db 841 TGGAGTGGGCGACCCCTACTTTCAGCTGCCCCCTGGAGAGAGCTTCGCGAAGTACACCG 900
QY 1058 CTTTCAACATCCCAAGCATCAACAGAGACCCCGCGCATCCGCTACAGTACACAGCTGC 1117
Db 901 CTTTCAACATCCCAAGCATCAACAGAGACCCCGCGCATCCGCTACAGTACACAGCTGC 960
QY 1118 TGCCCCAGGGCTGGAAGGCGACCCCGACATCTTCCAGAGCAGCATGACCAAGATCCCTGG 1177
Db 961 TGCCCCAGGGCTGGAAGGCGACCCCGACATCTTCCAGAGCAGCATGACCAAGATCCCTGG 1020
QY 1178 AGCCCTTCGCGCGCCGCAACCCCGAGATCTGATCTACAGTACATGAGACACCTGTACG 1237
Db 1021 AGCCCTTCGCGCGCGCCGCAACCCCGAGATCTGATCTACAGTACATGAGACACCTGTACG 1080
QY 1238 TGGGAGGAGCTGGAGATCGGCGAGCACCGCGCAAGATCGAGAGCTGCGCAAGCACCC 1297
Db 1081 TGGGAGGAGCTGGAGATCGGCGAGCACCGCGCAAGATCGAGAGCTGCGCGAGCACCC 1140
QY 1298 TGCTTGGCTGGGGCTTACACACCCCGCAACAGACACCAAGAGAGAGCCCTTCCTGT 1357
Db 1141 TGCTTGGCTGGGGCTTACACACCCCGCAACAGACACCAAGAGAGAGCCCTTCCTGT 1200
QY 1358 GGATGGGCTACGAGTGCACCCCGACAAAGTGSACCGTGCAGCCCATCAGCTGCCCGAGA 1417
Db 1201 GGATGGGCTACGAGTGCACCCCGACAAAGTGSACCGTGCAGCCCATCAGCTGCCCGAGA 1260
QY 1418 AGGAGAGCTGGACCTGAACACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCC 1477
Db 1261 AGGAGAGCTGGACCTGAACACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCC 1320
QY 1478 AGATCTACCCCGGATCAAGTGGCCAGCTGTGCAAGCTGCTGCGGCGCCCAAGGCCC 1537
Db 1321 AGATCTACCCCGGATCAAGTGGCCAGCTGTGCAAGCTGCTGCGGCGCCCAAGGCCC 1380
QY 1538 TGACCGACATCGTGCCTGTGACCGAGGAGCGGAGCTGGAGCTGCCCGAGAACCGCAGA 1597
Db 1381 TGACCGAGGTGATCCCTGTACCGAGAGGCGCGAGCTGGAGCTGGCGAGAACCGCAGA 1440
QY 1598 TCCTTCGCGAGCGCTGCACCGCTGTACTACGACCCCGACCAAGAGCTGTGTGCGCCAGA 1657
Db 1441 TCCTTCGCGAGCGCTGCACCGCTGTACTACGACCCCGACCAAGAGCTGTGTGCGCCAGA 1500
QY 1658 TCCAGAGAGCGGCGCAGACAGTGGACCTTACAGATCTTACAGAGCCCTTCAAGAAC 1717
Db 1501 TCCAGAGAGCGGCGCAGGCGCAGTGGACCTTACAGATCTTACAGAGCCCTTCAAGAAC 1560
QY 1718 TGAAGACCGCAAGTACGCAAGATCGCACCGCCCGACACCAAGAGCTGAAGCAGCTGA 1777
Db 1561 TGAAGACCGCAAGTACGCGCATGCGCGCGCCCGACACCAAGAGCTGAAGCAGCTGA 1620

QY 1778 CCGAGGCGCTGCAGAAAGATCGCCATGGAGAGCATCGTGTATCTGGGCAAGACCCCAAGT 1837
Db 1621 CCGAGGCGCTGCAGAAAGTGCAGCACCGAGAGCATCGTGTATCTGGGCAAGATCCCAAGT 1680
QY 1838 TCCGCTCGCCATCCAGAAAGAGACCTGGAGACCTGGTGGACCGACTACTGGCAGGCCA 1897
Db 1681 TCAAGCTGCCATCCAGAAAGAGACCTGGAGGCTTGGTGGATGGAGTACTGGCAGGCCA 1740
QY 1898 CTTGATCCCGAGTGGGAGTTCTGTGAACACCCCGCCCTGGTGAAGCTGTGTATCCAGC 1957
Db 1741 CTTGATCCCGAGTGGGAGTTCTGTGAACACCCCGCCCTGGTGAAGCTGTGTATCCAGC 1800
QY 1958 TGGAGAGGAGCCCATCATCGCGCGGCGAGACCTTCTACGTGAGACGCGCGCCCAACCGCG 2017
Db 1801 TGGAGAGGAGCCCATCGTGGCGCGGCGAGACCTTCTACGTGAGACGCGCGCGCCCAACCGCG 1860
QY 2018 AGACCAAGATCCGCAAGGCGGCTAGTGAACGACCGCGCGCGGCAAGAGATCGTGAGCC 2077
Db 1861 AGACCAAGCTGGCAAGGCGGCTAGTGAACGACCGCGCGCGGCGGCAAGAGTGGTGAGCA 1920
QY 2078 TGACCGAGACCAACCAAGAGACCGAGCTGCAGGCCATCCAGCTTGGCCCTTGACGAGCA 2137
Db 1921 TCGCCGACACCAACCAAGAGACCGAGCTGCAGGCCATCCACCTTGGCCCTTGACGAGCA 1980
QY 2138 GCGGAGCGAGGTGAACATCGTGCAGCAGCAGCAGTACGCCCTGGGCATCATCCAGGCC 2197
Db 1981 GCGGCTGGAGGTGAACATCGTGCAGCAGCAGCAGTACGCCCTGGGCATCATCCAGGCC 2040
QY 2198 AGCCCGACAAGAGCGAGAGCGAGCTGGTGAACGACATCATCAGCAGCTGATCAAGAGG 2257
Db 2041 AGCCCGACAAGAGCGAGAGCGAGCTGGTGAACGACATCATCAGCAGCTGATCAAGAGG 2100
QY 2258 AGAAGTGTACCTGAGCTGGTGCCTCCCGCCCAAGGGCATCGCGCGCAACAGCAGCAGCTG 2317
Db 2101 AGAAGTGTACCTGAGCTGGTGCCTCCCGCCCAAGGGCATCGCGCGCAACAGCAGCAGTGG 2160
QY 2318 ACAAGCTGTGAGCAAGGCGCATCCGCAAGGTGCTGTCTTCTGACGCGCATCGATGGCGCA 2377
Db 2161 ACAAGCTGTGAGCGCCGCGCATCCGCAAGGTGCTGTCTTCTGACGCGCATCGATGGCGCA 2220
QY 2378 TCGTATCTACAGTACATGGACGACCTGTAGTGGGAGCGCGCGCCCTAGGATCGATT 2437
Db 2221 TCGTATCTACAGTACATGGACGACCTGTAGTGGGAGCGCGCGCCCTAGGATCGATT 2280
QY 2438 AAAAGCTTCCCGGGCTAGCACCCTGGAATTC 2469
Db 2281 AAAAGCTTCCCGGGCTAGCACCCTGGAATTC 2312

RESULT 5

ABK91616

ID ABK91616 standard; DNA; 9166 BP.

XX ABK91616;

XX AC

XX XX

DT 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #168.

-DE

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

OS

XX

PN WO200232943-A2.

XX

PD 25-APR-2002.

XX

PF 14-AUG-2001; 2001WO-US25721.

XX

PR 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRAHARTI B K.
XX
PI Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
DR
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure; Page 766-769; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9166 BP; 2227 A; 2662 C; 2513 G; 1764 T; 0 other;

Query Match 83.1%; Score 2051.2; DB 24; Length 9166;
Best Local Similarity 92.6%; Pred. No. 1-2e-248;
Matches 2192; Conservative 0; Mismatches 163; Indels 13; Gaps 3;

QY 14 TGCGCCGAGCCATGAGCCAGGCGCACCGCCCAACATCTCTGATGCGAGGCGCATCGGACCCGTCG 3559
DB 2961 TGCGCCGAGCCATGAGCCAGGCGCACCGCCCAACATCTCTGATGCGAGGCGCATCGGACCCGTCG 361
QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGCGGCACATCGCCGCA 133
DB 3021 AGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGCGGCACATCGCCGCA 3080
QY 134 ACTGCCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGG 193
DB 3081 ACTGCCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGCAGGCGGCGGCGGCGGCGGCGGCGGCGG 3140
QY 194 AGGACTGACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 253
DB 3141 AGGACTGACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3199
QY 254 AGGCGCGCGAGTTCCCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 313
DB 3200 AGGCGCGCGAGTTCCCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3259
QY 314 AGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 364
DB 3260 AGGTGCGCGGAGAGAACAACTCCCTCTCAGAAGCAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3319
QY 365 ---CCCTGAACTTCCCGCAGCATCACTCTGTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
DB 3320 TATCTCTTTAGCTTCCCTTCAGATCACTCTTTGGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3379
QY 422 GCGGCCAGATCAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481
DB 3380 GGGCCAGCTGAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3439
QY 482 TGAGCCTTCCCGCGCAAGTGAAGCCCAAGATGATFCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
DB 3440 TGAACCTTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3499
QY 542 TGCGCCAGTACGACCATCTGATTCAGATCTCGGCGCAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601

DB 3500 TGCGCCGAGTACGACCCAGATCTCTGATCGAGATCTGCGGCCACAGGCGCATCGGACCCGTCG 3559
QY 602 TGATCGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGCA 661
DB 3560 TGGTGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGATCGGCTGCA 3619
QY 662 CCCTGAATTTCCCATCAGCCCATCGAGCCGTCGCGCGTGAAGCTGAAGCCCGCGCATGG 721
DB 3620 CCCTGAATTTCCCATCAGCCCATCGAGCCGTCGCGCGTGAAGCTGAAGCCCGCGCATGG 3679
QY 722 ACAGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCCCA 781
DB 3680 ACAGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGTCGAGA 3739
QY 782 TCTGCGAGGAGATGGAGAGGCGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 841
DB 3740 TCTGCGAGGAGATGGAGAGGCGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 3799
QY 842 ACACCCCGCTGTTGCGCATCAAGAAAGAGGACAGCACCAAGTGGCGGCAAGCTGGTGACT 901
DB 3800 ACACCCCGCTGTTGCGCATCAAGAAAGAGGACAGCACCAAGTGGCGGCAAGCTGGTGACT 3859
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCGATCCCCACC 961
DB 3860 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCGATCCCCACC 3919
QY 962 CCGCCGCGCTGAAGAAGAGAGCGTGCAGCGTGTGGAGCTGGGCGGCGGCGGCGGCGGCGGCGG 1021
DB 3920 CCGCCGCGCTGAAGAAGAGAGCGTGCAGCGTGTGGAGCTGGGCGGCGGCGGCGGCGGCGGCGG 3979
QY 1022 GCCTGCCCTTGGAGCGAGCTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAACA 1081
DB 3980 GCCTGCCCTTGGAGCGAGCTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAACA 4039
QY 1082 ACAGAGCCCGCGGATCCGCTACCACTACAAGTGTGCCCCAGGCGTGGAGGCGGCGGCGGCGG 1141
DB 4040 ACAGAGCCCGCGGATCCGCTACCACTACAAGTGTGCCCCAGGCGTGGAGGCGGCGGCGGCGG 4099
QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCGCAACCCCG 1201
DB 4100 CCGCCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCAAGCAACCCCG 4159
QY 1202 AGATGCTGATCTACCATGATGAGCAGCTGTACGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1261
DB 4160 ACATGCTGATCTACCATGATGAGCAGCTGTACGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4219
QY 1262 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCGTTCACCAACC 1321
DB 4220 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCGTTCACCAACC 4279
QY 1322 CCGACAAAGAGCAGCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1381
DB 4280 CCGACAAAGAGCAGCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4339
QY 1382 ACAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1441
DB 4340 ACAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4399
QY 1442 TCACAGAGCTGGTGGGCAAGCTGAACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1501
DB 4400 TCACAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4459
QY 1502 GCCAGCTGTCAAGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1561
DB 4460 GCCAGCTGTCAAGCTGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 4519
QY 1562 AGGAGCGCGGAGCTGGAGCTGGCCCGAGAACCGCGAGATCTCTGCGGCGGCGGCGGCGGCGGCGGCGG 1621
DB 4520 AGGAGCGCGGAGCTGGAGCTGGCCCGAGAACCGCGAGATCTCTGAGGAGGCGGCGGCGGCGGCGGCGG 4579
QY 1622 TGTACTACGACCCCGAGCAAGGAGCTGCTGGCGGAGATTCAGAAAGCAGGCGGCGGCGGCGGCGGCGGCGG 1681
DB 4580 TGTACTACGACCCCGAGCAAGGAGCTGATCGCGCGGAGATTCAGAAAGCAGGCGGCGGCGGCGGCGGCGG 4639

Db 3495 TGGCCAGTACGACAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCGTGC 3554
QY 602 TGATGCGCCCAACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGCA 661
Db 3555 TGGTGGCCCAACCCCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCTGCA 3614
QY 662 CCCTGAATTCCTCCATAGCCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGATGG 721
Db 3615 CCCTGAATTCCTCCATAGCCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGATGG 3674
QY 722 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACGAGAGAGATCAAGGCCCTGACGCCCA 781
Db 3675 ACGGCCCCAAGGTGAAGCAGTGGCCCTGACGAGAGAGATCAAGGCCCTGCTGGAGA 3734
QY 782 TCTCGAGGAGATGGAGAAGGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACA 841
Db 3735 TCTGCACGAGATGGAGAAGGGGCAAGATCAGCAAGATCGGCCCGGAGAACCCCTACA 3794
QY 842 ACACCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT 901
Db 3795 ACACCCCGTGTTCGCCATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACT 3854
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCCCACC 961
Db 3855 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCCCACC 3914
QY 962 CCGCCGCTTGAAGAAGAGAGGCTGACCGTGTGGACGTGGCGAGCGCTACTTCA 1021
Db 3915 CCGCCGCTTGAAGCAGAAGAGCGTGACCGTGTGGACGTGGCGAGCGCTACTTCA 3974
QY 1022 GCGTGGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTTCACCATCCCGCAGATCAACA 1081
Db 3975 GCGTGGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTTCACCATCCCGCAGATCAACA 4034
QY 1082 ACAGACCCCGCGCATCCGCTACAGTACAAAGCTGTGCCAGAGGCTGGAAGGGCAGCC 1141
Db 4035 ACAGACCCCGCGCATCCGCTACAGTACAAAGCTGTGCCAGAGGCTGGAAGGGCAGCC 4094
QY 1142 CCAGCATCTTCCAGACAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAAGCCCG 1201
Db 4095 CCAGCATCTTCCAGTACAGCATGACCAAGATCCTGGAGCCCTTCCGCGCAAGCAGACCCCG 4154
QY 1202 AGATCTGATCTACAGTACATGGACGACCTGTACGTGGCGCAGCAGCTGGAGATCGGCC 1261
Db 4155 ACATCTGATCTACAGTACATGGACGACCTGTACGTGGCGCAGCAGCTGGAGATCGGCC 4214
QY 1262 AGCACCCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGCTGGGCTTCACCAACC 1321
Db 4215 AGCACCCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGCTGGGCTTCACCAACC 4274
QY 1322 CCGACAAGACCAAGAGGAGCCCGCTTCTGTGGATGGGCTACGAGCTGCACCCCG 1381
Db 4275 CCGACAAGAGCACCAGAAGGAGCCCGCTTCTGTGGATGGGCTACGAGCTGCACCCCG 4334
QY 1382 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGAGAGCTGGACCGTGAACGACA 1441
Db 4335 ACAAGTGGACCGTGCAGCCCATCGTCTGCCCGAGAGAGACAGCTGGACCGTGAACGACA 4394
QY 1442 TCAGAAGCTGTGGCAAGCTGAAGTGGCGCAGCCAGATCTACCCCGCATCAAGGTGC 1501
Db 4395 TCAGAAGCTGTGGCAAGCTGAAGTGGCGCAGCCAGATCTACCCCGCATCAAGGTGC 4454
QY 1502 GCGAGTGTGAAGTGTGTCGCGCGCGCCCAAGGCCCTGACGACATCTGTCGCCCTGACCG 1561
Db 4455 GCGAGTGTGAAGTGTGTCGCGCGCGCCCAAGGCCCTGACGAGTGTGTCGCCCTGACCG 4514
QY 1562 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCCGAGCCCGTGCACGGCG 1621
Db 4515 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGAGGAGCGCGTGCACGGCG 4574
QY 1622 TGTACTACGACCCAGCAAGGACCTGGTGGCGGAGATCCAGAGCAGGAGGCGGACGACCT 1681
Db 4575 TGTACTACGACCCAGCAAGGACCTGATCGCCGAGATCCAGAGCAGGAGGCGGACGACCT 4634

RESULT 7

ABK91617

ID ABK91617 standard; DNA; 9169 BP.

XX ABK91617;

XX ABK91617;

XX 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #169.

DE HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

XX WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.
 PR 28-MAR-2001; 2001US-279257P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (CHAD/) CHADABARTI B K.
 XX
 PI Nabel GJ, Huang Y;
 XX WPI; 2002-452382/48.
 DR
 XX
 PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
 PT its encoded protein, useful as vaccines for genetic or human
 PT immunisation for acquired immunodeficiency syndrome or human
 PT immunodeficiency virus infection -
 XX
 XX Disclosure; Page 769-772; 794pp; English.
 PS
 XX
 CC The invention relates to a nucleic acid molecule encoding a modified HIV
 CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
 CC genetic or protein immunisation to a host, respectively. In particular
 CC these are useful for ameliorating the symptoms of acquired
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
 CC compositions are useful for treating or preventing HIV infections or
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
 CC plasmid DNA sequences of the invention.
 XX
 SQ Sequence 9169 BP; 2227 A; 2668 C; 2505 G; 1769 T; 0 other;

Query Match 82.2%; Score 2028.8; DB 24; Length 9169;
 Best Local Similarity 91.7%; Pred. No. 7.5e-246;
 Matches 2194; Conservative 0; Mismatches 182; Indels 16; Gaps 4;

Db 3503 TGGGCCAGTACGACACAGATCCCTGATCGAGATCTGCGGCCCAAGGCCATCGGCACCGTGC 3562
 QY 602 TGATCGGCCCCACCCCGGTGAACATCATCGGCCGCAACATGCTGACCCAGCTGGGTGCA 661
 Db 3563 TGGTGGGCCCCACCCCGGTGAACATCATCGGCCGCAACCTGCTGACCCAGATCGGCTGCA 3622
 QY 662 CCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTCGTGAAGCTGAAGCCCGCATGG 721
 Db 3623 CCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTCGTGAAGCTGAAGCCCGCATGG 3682
 QY 722 AGGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCA 781
 Db 3683 AGGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGGTGGAGA 3742
 QY 782 TCTGGAGAGATGAGAGAGGAGGCAAGATCACAAGATCGGCCCGGAGAACCCCTACA 841
 Db 3743 TCTGCACCGAGATGAGAGAGGAGGCAAGATCAGCAAGATCGGCCCGGAGAACCCCTACA 3802
 QY 842 ACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAGTGGCGCAAGCTGGTGACT 901
 Db 3803 ACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAGTGGCGCAAGCTGGTGACT 3862
 QY 902 TCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCCACC 961
 Db 3863 TCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCCACC 3922
 QY 962 CGCGCGGCTGAAGAAGAAGAGAGCGTGAACGCTGCTGAGACGTGGCGAGCGCTACTTCA 1021
 Db 3923 CGCGCGGCTGAAGAAGAAGAGAGCGTGAACGCTGCTGAGACGTGGCGAGCGCTACTTCA 3982
 QY 1022 GCGTCCCGCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGACATCAACA 1081
 Db 3983 GCGTCCCGCTGGACAGAGACTTCCGCAAGTACACCGCTTCCACCATCCCGACATCAACA 4042
 QY 1082 ACGAGACCCCGGCATCCGCTACCACTACAAGTCTGCCCGAGGCTGGAGAGCGAGCC 1141
 Db 4043 ACGAGACCCCGGCATCCGCTACCACTACAAGTCTGCCCGAGGCTGGAGAGCGAGCC 4102
 QY 1142 CCAGCATCTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCCGACACCCGG 1201
 Db 4103 CGGCATCTTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCCGACACCCGG 4162
 QY 1202 AGATCGTATCTACCACTACATGGAGACGCTTACCTGGGAGCGACCTTGGAGATCGGCC 1261
 Db 4163 ACATCGTATCTACCACTACATGGAGACGCTTACCTGGGAGCGACCTTGGAGATCGGCC 4222
 QY 1262 AGCAGCGCGCAAGATCGAGGAGTGCAGACCTGCTGCGCTGGGCTTCCACACCC 1321
 Db 4223 AGCAGCGCGCAAGATCGAGGAGTGCAGACCTGCTGCGCTGGGCTTCCACACCC 4282
 QY 1322 CCGACAAGACGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1381
 Db 4283 CCGACAAGACGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4342
 QY 1382 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1441
 Db 4343 ACAAGTGGACCGTGCAGCCCATCGTGTGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 4402
 QY 1442 TCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAGGTGC 1501
 Db 4403 TCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAGGTGC 4462
 QY 1502 GCCAGCTGCAAGCTGCTGCGCGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1561
 Db 4463 GCCAGCTGCAAGCTGCTGCGCGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4522
 QY 1562 AGGAGCGGAGCTGGAGCTGGCCCGAGAACCCGAGATCTCTCGGAGGAGGAGGAGGAGGAGGAG 1621
 Db 4523 AGGAGCGGAGCTGGAGCTGGCCCGAGAACCCGAGATCTCTCGGAGGAGGAGGAGGAGGAGGAG 4582
 QY 1622 TGTACTACGACCCAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1681
 Db 4583 TGTACTACGACCCAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4642

```
QY 1682 GGACCTACACAGATCTACAGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCAAGA 1741
DB 4643 GGACCTACACAGATCTACAGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCGCA 4702
QY 1742 TGGCGACCGCCACACACAGAGCTGAAGCAGCTGACCGAGGCGCGTGAGAAAGATCGCCA 1801
DB 4703 TGAAGGGCGCCACACACAGAGCTGAAGCAGCTGACCGAGGCGCGTGAGAAAGATCGCCA 4762
QY 1802 TGAAGAGCATCTGATGCTGGGGCAAGACCCCAAGTTCCCGCTGCCATCCACAAGAGA 1861
DB 4763 CCAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCATCCACAAGAGA 4822
QY 1862 CTTGGAGACCTCTGAGACCGACTACTGGCAGCGCCACTGGATCCCGAGTGGAGTTCG 1921
DB 4823 CTTGGAGGCTCTGAGACCGAGTACTGGCAGCGCCACTGGATCCCGAGTGGAGTTCG 4882
QY 1922 TGAACACCCCCCTGGTGAAGCTGTGTGACAGCTGGAGAGGAGCCCATCATCGGCG 1981
DB 4883 TGAACACCCCCCTGGTGAAGCTGTGTGACAGCTGGAGAGGAGCCCATCATCGGCG 4942
QY 1982 CCGAGACCTTCTAGTGGAGCGCGCCGCCAACCGCGAGACCAAGATCGGCAAGCGCGCT 2041
DB 4943 CCGAGACCTTCTAGTGGAGCGCGCGCCGCCAACCGCGAGACCAAGATCGGCAAGCGCGCT 5002
QY 2042 ACCTGACCGACCGGGCGCGCAGAGATCTGTGACCTGTGACCGAGACCAACCAAGAGA 2101
DB 5003 ACCTGACCGACCGGGCGCGCAGAGATCTGTGACCTGTGACCGAGACCAACCAAGAGA 5062
QY 2102 CCGAGCTGACAGGCCATCCAGCTGGCCCTGTGAGGACAGCGCGAGAGTGAACATCGTGA 2161
DB 5063 CCGAGCTGACAGGCCATCCAGCTGGCCCTGTGAGGACAGCGCGAGAGTGAACATCGTGA 5122
QY 2162 CCGACACCGAGTACCGCTGGGCATCATCCAGGCCCGCCAGACAGGAGGAGCGAGC 2221
DB 5123 CCGACACCGAGTACCGCTGGGCATCATCCAGGCCCGCCAGACAGGAGGAGCGAGC 5182
QY 2222 TGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGGTGC 2281
DB 5183 TGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGGTGC 5242
QY 2282 CCGCCCAACAGGCGATCGCGGCGCAACGAGCAGATCGACAGCTGGTGAGCAAGGGCATCC 2341
DB 5243 CCGCCCAACAGGCGATCGCGGCGCAACGAGCAGGTGGACGGCCTGGTGAGCGCGGCATCC 5302
QY 2342 GCAAGGTGCTTCTCTGAGCGGATCGATGGCGCATCTGATCTACCAAGTA 2393
DB 5303 GCAAGGTGCTTCTCTGAGCGGATCGATGGCGCATCTGATCTACCAAGTA 5354

RESULT 8
ABK91611
ID ABK91611 standard; DNA; 9194 BP.
XX
AC ABK91611;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #163.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
XX Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN WO200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001WO-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
```

```
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRAPARTI B K.
XX
XX Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
```

```
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
XX its encoded protein, useful as vaccines for genetic or protein
XX immunisation for acquired immunodeficiency syndrome or human
XX immunodeficiency virus infection
XX
XX Disclosure; Page 750-753; 794pp; English.
XX
```

```
XX The invention relates to a nucleic acid molecule encoding a modified HIV
XX (human immunodeficiency virus) protein. The nucleic acid molecule and its
XX encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
XX genetic or protein immunisation to a host, respectively. In particular
XX these are useful for ameliorating the symptoms of acquired
XX immunodeficiency syndrome (AIDS) or HIV infection and generating an
XX antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
XX Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
XX compositions are useful for treating or preventing HIV infections or
XX AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
XX plasmid DNA sequences of the invention.
XX
```

XX Sequence 9194 BP; 2232 A; 2671 C; 2524 G; 1767 T; 0 other;

Query Match 82.1%; Score 2027; DB 24; Length 9194;
Best Local Similarity 91.6%; Pred. No. 1.3e-245;
Matches 2194; Conservative 0; Mismatches 185; Indels 16; Gaps 4;

```
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACGAGCGCCACATCTCTGATCAGCGCGGCAACT 70
DB 2967 TGGCCGAGGCGCATGAGCCAGGTTGACCAACGCGCCACCATCATGATCAGCGCGGCAACT 3026
QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGGCCACATCGGCC 130
DB 3027 TCCGCAACGAGCGCAAGATCGTGAAGTGTCTCAACTGCGCAAGGAGGGCCACACGCC 3086
QY 131 GCAACTCCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 190
DB 3087 GCAACTCCGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGCACGAGCGAGCGCCAGCCAACTTCTTCGCGAGGAGCTGGCCCTTCCCGCAGG 250
DB 3147 TGAAGGACTGCACGAGCGAGCGAGCGCTTA-TTTTATAGGAAGATCTGGCCCTTCCACAG 3205
QY 251 GCAAGGCGCGGAGTTCCCGAGGAGCAGAACCGCGCCCAACAGCCCGCCACCGCGGAGC 310
DB 3206 GGAAGGCGCGGAGTTTCTTCAGAGCAGACACAGAGCCACAGCCCGCCACAGAGAGAGC 3265
QY 311 TGCAGGTGCGGGG-----CGACAACCCCGCAGGAGCGCGCGCGCGAGCGCGCGGCA 364
DB 3266 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAGAGGAGAGCGGATAGACAAGGAA 3325
QY 365 -----CCCTGAACCTCCCGCAGATCACCTGTGCGAGCGCCCTGCTGAGCATCAAGG 418
DB 3326 CTGTATCTCTTAGCTTCCCTCAGATCACTTTTGGCAGCAGACCCCTCTGTCACATTAAGA 3385
QY 419 TGGGCGCGCGAGATCAAGGAGCGCCCTGCTGGACACCGCGCGCGAGCAGACACCGTGTGGAGG 478
DB 3386 TAGGGGCGCGAGTGAAGGAGCGCCCTTCTAGACACCGCGCGCGAGCAGACACCGTGTGGAGG 3445
QY 479 AGATGAGCCTCCCGCGCAAGTGGAGCCCAAGATGATCGCGGGCATCGCGGGCTTCATCA 538
DB 3446 AGATCAACCTCCCGCGCGCTTGGAGGCCCAAGATGATCGCGGGCATCGCGGGCTTCATCA 3505
QY 539 AGGTGGCGCGAGTACGACGAGATCTGTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
```

Db 3506 AGGTGGGCGAGTACGACCCAGATCCTGATCGAGATCTCGGSCCAAGAGCCATCGGCACCG 3565
QY 599 TGTGTATCGGCCCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGCT 658
Db 3566 TGTGTGTGGGCCCCACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGCT 3625
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCCCATCGAGACCCGTGAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTGAACTTCCCATCAGCCCATCGAGACCCCATCGAGACCCGTGAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGGCCCGGAGTGAAGCAAGTGGGCAAGGATACCAAGATCGGCCCGCGAAGCCCT 838
Db 3686 TGGACGGCCCGGAGTGAAGCAAGTGGGCAAGGATACCAAGATCGGCCCGCGAAGCCCT 3805
QY 779 CCATCTGCGAGGAGTGAAGCAAGGATGAGGCAAGGATACCAAGATCGGCCCGCGAAGCCCT 838
Db 3746 AGATCTGCAAGGATGAGGCAAGGATGAGGCAAGGATACCAAGATCGGCCCGCGAAGCCCT 3805
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAGAGACAGCACCAAGTGGCGCAAGCTGGTG 898
Db 3806 ACAACACCCCGTGTTCGCCATCAAGAAGAGACAGCACCAAGTGGCGCAAGCTGGTG 3865
QY 899 ACTTCGCGAGTGAACAGCGCACCCAGACTTCTGGAGGTGAGCTGGGCATCCCC 958
Db 3866 ACTTCGCGAGTGAACAGCGCACCCAGACTTCTGGAGGTGAGCTGGGCATCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAGAGAGCGTGCACCGTGTGACGTTGGCGACCGCTACT 1018
Db 3926 ACCCGCGCGCTGAAGAAGAGAGCGTGCACCGTGTGACGTTGGCGACCGCTACT 3985
QY 1019 TCAGGTGCCCCGTGACGAGGACTTCCGCAAGTACACCGCTTACCATTCCCGAGATCA 1078
Db 3986 TCAGGTGCCCCGTGACGAGGACTTCCGCAAGTACACCGCTTACCATTCCCGAGATCA 4045
QY 1079 ACAACGAGACCCCGGCTACCGTACCAAGTACCAAGTCTGGAGCCCTTCCGCGCCGCAAC 1198
Db 4046 ACAACGAGACCCCGGCTACCGTACCAAGTACCAAGTCTGGAGCCCTTCCGCGCCGCAAC 4105
QY 1139 GCGCGAGATCTTCAGAGCAGCATGACCAAGATCGTGGAGCCCTTCCGCGCCGCAAC 1198
Db 4106 GCGCGAGATCTTCAGTGCAGCATGACCAAGATCGTGGAGCCCTTCCGCGAGCAGAC 4165
QY 1199 CCGAGATCGTATCTACAGTATACATGGAGACCTGAGTGGCGAGCGACCTGGAGATCG 1258
Db 4166 CCGAGATCGTATCTACAGTATACATGGAGACCTGAGTGGCGAGCGACCTGGAGATCG 4225
QY 1259 GCGAGACCGCGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 1318
Db 4226 GCGAGACCGCGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 4285
QY 1319 CCGCGACAAGACCAAGAGGAGCCCGCTTCTGTGGATGGGCTACGAGCTGCACC 1378
Db 4286 CCGCGACAAGAGCAGAGGAGGAGCCCGCTTCTGTGGATGGGCTACGAGCTGCACC 4345
QY 1379 CCGACAAGTGGAGCGTGCAGCCATCGAGCTGCGCGAAGAGAGAGCTGGACCGTGAACG 1438
Db 4346 CCGACAAGTGGAGCGTGCAGCCATCGTGGCTGCGCGAAGAGAGAGCTGGACCGTGAACG 4405
QY 1439 ACATCCAGAAGCTGGTGGGCAAGTGAATGGCGAGCGAGATACCGCGGCGATCAAGG 1498
Db 4406 ACATCCAGAAGCTGGTGGGCAAGTGAATGGCGAGCGAGATACCGCGGCGATCAAGG 4465
QY 1499 TGGCGAGCTGTGAAGTGTGCGCGCGCGCAAGGCGCTGACCGACATCTGCGCCCTGA 1558
Db 4466 TGGCGAGCTGTGAAGTGTGCGCGCGCGCAAGGCGCTGACCGAGGTGGTGGCCCTGA 4525
QY 1559 CCGAGGAGCGGAGCTGGAGCTGCGCGAGAACCGGAGATCTTGGCGAGCGCCGTGCACG 1618
Db 4526 CCGAGGAGCGGAGCTGGAGCTGCGCGAGAACCGGAGATCTTGAAGGAGCGCCGTGCACG 4585
QY 1619 GCGTGTACTAGACCCCGAGCAAGGACCTGGTGGCGGAGATCCAGAAGCAGGCGCACGACC 1678
Db 4586 GCGTGTACTAGACCCCGAGCAAGGACCTGTGTCGCGGAGATCCAGAAGCAGGCGCACGACC 4645

RESULT 9

ABK91613

ID ABK91613 standard; DNA; 12411 BP.

XX AC ABK91613;

XX DT 14-AUG-2002 (first entry)

XX DE Modified HIV protein-encoding plasmid DNA #165.

XX DE HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

XX WO200232943-A2.

XX PD 25-APR-2002.

XX PF 14-AUG-2001; 2001WO-US25721.

XX PR 14-AUG-2000; 2000US-225097P.

PR	14-NOV-2000; 2000US-25211SP.
PR	28-MAR-2001; 2001US-279257P.
XX	(US\$) US DEPT HEALTH & HUMAN SERVICES.
PA	(CHAD/) CHADRABARTI B K.
XX	Nabel GJ, Huang Y;
PI	WPI; 2002-452382/48.
XX	New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and its encoded protein, useful as vaccines for genetic or protein immunisation for acquired immunodeficiency syndrome or human immunodeficiency virus infection -
XX	Disclosure; Page 756-760; 794pp; English.
XX	The invention relates to a nucleic acid molecule encoding a modified HIV (human immunodeficiency virus) protein. The nucleic acid molecule and its encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for genetic or protein immunisation to a host, respectively. In particular these are useful for ameliorating the symptoms of acquired immunodeficiency syndrome (AIDS) or HIV infection and generating an antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env, Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent compositions are useful for treating or preventing HIV infections or AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding plasmid DNA sequences of the invention.
XX	Sequence 12411 BP; 2938 A; 3733 C; 3476 G; 2264 T; 0 other;
SQ	Query Watch 82.1%; Score 2027; DB 24; Length 12411; Best Local Similarity 91.6%; Pred. No. 1.2e-245; Matches 2194; Conservative 0; Mismatches 185; Indels 16; Gaps 4
QY	14 TGCGCGAGGCCATGAGCCAGG---CCACAGGCGCAACATCCTCATGCAGCGAGCAACT 70
Db	2967 TGCGCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATTATGATGACGCGGGCACT 3026
QY	71 TCAAGGGCCCCAAGCGCATCATCAAGTGGCTTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
Db	3027 TC CGCAACACAGCGCAGATGCTGAAGTGCTTCAACTGCGGCAAGGAGGGCCACACGCC 3086
QY	131 GCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTCGCGCAAGAGGGCCACACAGA 190
Db	3087 GCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTCGCGCAAGAGGGCCACACAGA 3146
QY	191 TGAAGACTGCAACCGAGCGCCAGGCACTTCTCCGCGAGGACCTGGCCTTCCCGCAGG 250
Db	3147 TGAAGACTGCAACCGAGCGCACAGGCTAA-TTTT'TTAGGGAAGATCTGGCCTTCCCAAG 3205
QY	251 GCAAGGCCGCGAGTTTCCCGAGGAGCAACCGCGCCCAACAGCCGCCACACAGCCGCGAGC 310
Db	3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACACAGCCCGCAGAGAGAGC 3265
QY	311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCCAGCGCCAGGCA 364
Db	3266 TTCAGGTTTGGGAAGAGACAACAACCTCCTCTCAGAAGCAGGAGCCGATAGACAAGGAA 3325
QY	365 -----CCCTGAACCTTCCCCAGATACCTCTGTGCAGCGCCCGCTGTGTGAGATCAAGG 418
Db	3326 CTGTATCTTTAGCTTCCCTTCAGATCACTCTTTGACGCGGACCCCTCTGTACATAAAGA 3385
QY	419 TGGCGGCGCCAGATCAAGAGGGCCCTGCTGGACACCGCGCGCCAGACACCGTGTGGAGG 478
Db	3386 TAGGGGCGCAGCTGAAGGAGGCCCTTCTAGACACCGGCGCGCCAGCACACCGTGTGGAGG 3445
QY	479 AGATGAGCTTGCCCGGCAAGTGAAGCCCAAAGATGATCGCGGGCATCGCGGCTTCATCA 538
Db	3446 AGATGAACCTTGCCCGGCGCGTGGAGGCCCAAAGATGATCGCGGGCATCGCGGCTTCATCA 3505
QY	539 AGGTGCGCCAGTAGCACGACAGATCCTGATCGAGATCTGGGGCAAGAAGGCCATCGGCACCG 598

Db 3506 AGGTGGCCAGTAGACACAGATCCTGATCGAGATCTGCGGCCACAAGGCCATCGGCACCG 3565
QY 599 TGTGTGATGGCCCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3566 TGTGTGATGGCCCCACCCCGTGAACATCATCGGCGCGCAACCTGCTGACCCAGATCGGCT 3625
QY 659 GCACCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCTGTAAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCTGTAAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGACGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGTGTG 3745
QY 779 CCATCTCGGAGGAGATGGAGAAGGAGGCAAGATCAACCAAGATCGGCCCGCGAGAACCCT 838
Db 3746 AGATCTGCACCGAGATGGAGAAGGAGGCAAGATCAAGAGATCGGCCCGCGAGAACCCT 3805
QY 839 ACAACACCCCGCTGTTGCGCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTG 898
Db 3806 ACAACACCCCGCTGTTGCGCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTG 3865
QY 899 ACTTCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGGATCCCC 958
Db 3866 ACTTCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGGATCCCC 3925
QY 959 ACCCCGCGGCTTGAAGAAGAAGAGGCTGACCGTGTGGAGCTGGGCGACGCCCTACT 1018
Db 3926 ACCCCGCGGCTTGAAGAAGAAGAGAGCGTGACCGTGTGGAGCTGGGCGACGCCCTACT 3985
QY 1019 TCAGCTGCCCCCTGGACGAGGACTTCCGAAGTACACCGCTTCAACCATCCCCAGCATCA 1078
Db 3986 TCAGCTGCCCCCTGGACGAGGACTTCCGAAGTACACCGCTTCAACCATCCCCAGCATCA 4045
QY 1079 ACACAGAGACCCCGCATCCGCTACAGTACAACTGTGCCCCAGGGCTGAAGGGCA 1138
Db 4046 ACACAGAGACCCCGCATCCGCTACAGTACAACTGTGCCCCAGGGCTGAAGGGCA 4105
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCGCCGCACTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAACC 4165
QY 1199 CCGAGATCGTGATACCAAGTACATGGACACCTGTACGTGGGCGACGACCTGGAGATCG 1258
Db 4166 CCGACATCGTGATACCAAGTACATGGACACCTGTACGTGGGCGACGACCTGGAGATCG 4225
QY 1259 GCGACACCGCGCAAGATCGAGGAGCTGGCAAGCAGCTGCTGCGCTGGGGCTTCAACA 1318
Db 4226 GCGACACCGCGCAAGATCGAGGAGCTGGCGAGCAGCTGCTGCGCTGGGGCTTCAACA 4285
QY 1319 CCGCCGACAGAAGACCAAGAGGAGCCCGCTTCCGTGGATGGGCTACGAGCTGCACC 1378
Db 4286 CCGCCGACAGAAGACCAAGAGGAGCCCGCTTCCGTGGATGGGCTACGAGCTGCACC 4345
QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGTGGCCCGAGAGGAGAGCTGGACCGTGAACG 1438
Db 4346 CCGACAAGTGGACCGTGCAGCCCATCGAGTGGCCCGAGAGGAGAGCTGGACCGTGAACG 4405
QY 1439 ACATCCAGAAGCTGGTGGCGCAAGCTGAATGGGCGAGCCAGATCTACCCCGGCATCAAG 1498
Db 4406 ACATCCAGAAGCTGGTGGCGCAAGCTGAATGGGCGAGCCAGATCTACCCCGGCATCAAG 4465
QY 1499 TCGCCAGCTGTCAAGCTGTCTGCGCGGCGCAAGGCCCTGACCGACATCTGTGCCCTGCA 1558
Db 4466 TCGCCAGCTGTCAAGCTGTCTGCGCGGCGCAAGGCCCTGACCGAGTGTGTGCCCTGCA 4525
QY 1559 CCGAGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGCGAGCCCGTGCAG 1618
Db 4526 CCGAGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGCGAGCCCGTGCAG 4585
QY 1619 GCGTGTACTACGACCCCGACAGGAGCTGCTGGCGGAGATCCAGAGAGAGGCGCAGACC 1678
Db 4586 GCGTGTACTACGACCCCGACAGGAGCTGATCGCGAGATCCAGAGAGGCGCAGGGCC 4645

QY 1679 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCA 1738
Db 4646 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCC 4705
QY 1739 AGATGGCGACCGCCCAACACAGCAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCG 1798
Db 4706 GCATGAAGGCGCCCAACACAGCAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCG 4765
QY 1799 CCATGAGAGCATCTGTATCTGGGCGAAGACCCCAAGTTCCCGCTGCCCATCAGAAG 1858
Db 4766 CCACGAGAGCATCTGTATCTGGGCGAAGACCCCAAGTTCAAGCTGCCATCCAGAAG 4825
QY 1859 AGACCTGGGAGACCTGTGTGAGCAGCTACTGTGCGAGGCCACTGTGATCCCGAGTGGGAGT 1918
Db 4826 AGACCTGGGAGGCGCTGTGTGAGCAGCTACTGTGCGAGGCCACTGTGATCCCGAGTGGGAGT 4885
QY 1919 TCGTGAACACCCCCCTCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCCATCATCG 1978
Db 4886 TCGTGAACACCCCCCTCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCCATCATCG 4945
QY 1979 GCGCCGAGAGCTTCTACGTGGAGCGGCGGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 2038
Db 4946 GCGCCGAGAGCTTCTACGTGGAGCGGCGGCCCAACCGCGAGACCAAGCTGGCAAGGCCG 5005
QY 2039 GCTACCTGACCGACCGGCGCGGCGCAGAAAGATCCTGAGCCCTGACCGAGACCAACCA 2098
Db 5006 GCTACCTGACCGACCGGCGCGGCGCAGAAAGTGTGTCCCTGACCGACACCAACCA 5065
QY 2099 AGACCGAGCTGCGAGGCGCATCCAGCTGGCCCTGCAGGACAGCGCAGAGGTGAACATCG 2158
Db 5066 AGACCGAGCTGCGAGGCGCATCCACCTGGCCCTGCAGGACAGCGGCTGGAGGTGAACATCG 5125
QY 2159 TGACCGACAGCCAGTAGTACGCCCTGGGCATCATTCAGGCCCGACCGCCGAGAGCGAGCG 2218
Db 5126 TGACCGACAGCCAGTAGTACGCCCTGGGCATCATTCAGGCCCGACCGCCGAGAGCGAGCG 5185
QY 2219 AGCTGTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2278
Db 5186 AGCTGTGTGAGCCAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 5245
QY 2279 TGCCCGCCCGACAGGCGATCGGCGGCAACGACAGCAGATCGCAAGAGCTGGTGAAGCAAGGCA 2338
Db 5246 TGCCCGCCCGACAGGCGATCGGCGGCAACGAGCAGGTGGACGCGCTGGTGAAGCGCGCA 5305
QY 2339 TCCGCAAGCTGTGTCTTCTGGAGCGCATCGATGGCGGCATCGTGTATCTACCACTA 2393
Db 5306 TCCGCAAGCTGTGTCTTCTGGAGCGCATCGAACAAGGCCCGCAGGAGGAGCAGAGAA 5360

RESULT 11

AAA70479

ID AAA70479 standard; DNA; 2306 BP.

XX

AC AAA70479;

XX

DT 28-NOV-2000 (first entry)

XX

DE HIV FS(-)_ProMod_RTopt_YM coding sequence.

XX

KW HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.

XX

OS Human immunodeficiency virus type 1.

XX

OS Synthetic.

PN WO200039302-A2.

XX

PD 06-JUL-2000.

XX

PF 30-DEC-1999; 99WO-US31245.

XX

PR 31-DEC-1998; 98US-0114495.

PR 01-DEC-1999; 99US-0168471.

(CHIR) CHIRON CORP.
 Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;
 Greer C, Selby M, Walker C;
 WPI; 2000-452400/39.
 Expression cassettes encoding the human immunodeficiency virus (HIV)
 Gag-containing polypeptide useful for vaccinating against HIV
 infections and acquired immunodeficiency syndrome (AIDS) -
 Claim 7; Fig 73; 391pp; English.
 The present sequence is the coding sequence of a synthetic HIV
 Gag-reverse transcriptase expression cassette, PS(-).ProtMod.Rtopt_YM.
 The Gag protein of HIV is needed for the assembly of virus-like
 particles. In addition, the Gag protein is involved in many stages of the
 HIV life cycle, including assembly, virion maturation after particle
 release and early post-entry steps in viral replication. The expression
 cassette may be used for the recombinant expression of HIV
 Gag-polypeptides which may then be used to vaccinate against HIV
 infection and acquired immunodeficiency syndrome (AIDS).
 Query Match 82.0%; Score 2025.2; DB 21; Length 2306;
 Best Local Similarity 93.3%; Pred. No. 2.6e-245;
 Matches 2156; Conservative 0; Mismatches 138; Indels 18; Gaps 3;
 QY 170 GCGCAAGAGGGCCACAGATGAAGAGTGCACCGAGCGCCAGGCCAACTTCTTCGCG 229
 DB 1 GCGCGCGGAAGGACACCAAAATGAAGATTGCATGAGAGACAGGCTAAATTCCTCGCG 60
 QY 230 AGNACCTGGCTTCCCGGAGGCGAGCGCGAGTTCGCCAGGAGACAGACCGCGCCA 289
 DB 61 AGNACCTGGCTTCCCGGAGGCGAGCGCGAGTTCAGCAGGAGACAGACCGCGCCA 120
 QY 290 ACAGCCCCACAGCGCGAGTGCAGTGGCGGCGG-----ACAAACCCCGCAGCGAGG 343
 DB 121 ACAGCCCCACCGCGGAGTGCAGTGGCGGCGGAGACAGCGCTGAGCGAGG 180
 QY 344 CCGCGCGCGGCGGCGGAGCGGACCGCTG-----AACTTCCCGGAGATCACCGTGTGGCAGC 397
 DB 181 CCGCGCGCGGCGGCGGAGCGGACCGTGAAGTCAACTTCCCGGAGATCACCGTGTGGCAGC 240
 QY 398 GCGCCCTGGTGAAGTCAAGTGGCGGCGGAGATCAAGGAGCGGCTGCTGACACCGCGG 457
 DB 241 GCGCCCTGGTGAAGTCAAGTGGCGGCGGAGTCAAGGAGCGGCTGCTGACACCGCGG 300
 QY 458 CCGACGACACCGTGTGGAGGAGATGAGCCTGCGCGGCGAAGTGAAGCCCAAGATGATCG 517
 DB 301 CCGACGACACCGTGTGGAGGAGATGAACCTGCGCGGCGAAGTGAAGCCCAAGATGATCG 360
 QY 518 GCGCGATCGCGGCTTCATCAAGTGGCGGCGGAGTACGACCGAGATCCTGATCGAGATCTGCG 577
 DB 361 GCGGGATCGGGGCTTCATCAAGTGGCGGCGGAGTACGACCGAGATCCCGGTGGAGATCTGCG 420
 QY 578 GCAGAGAGGCGATCGGACCGTGTGATCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCA 637
 DB 421 GCCACAGGCGATCGGACCGTGTGTTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGCA 480
 QY 638 ACATGCTGACCGAGTGGCTGCACTGACCTGAACTTCCCGATCAGCGCCATCAGACCGTGC 697
 DB 481 ACCTGCTGACCGAGTGGCTGCACTGACCTGAACTTCCCGATCAGCGCCATCAGACCGTGC 540
 QY 698 CCGTGAAGTGAAGCGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 757
 DB 541 CCGTGAAGTGAAGCGCGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
 QY 758 AGAAGATCAAGGCGCTGACCGCGCTTCCGAGGAGATGGAGAGGCGGCGGCGGCGGCGGCGG 817
 DB 601 AGAAGATCAAGGCGCTGTTGAGATCTGCACCGGAGATGGAGAGGCGGCGGCGGCGGCGGCGG 660

QY 818 AGATCGGCCCGGAGACCCCTACAAACACCCCGCTGTTCCGCCATCAAGAAGAAGACAGCA 877
 DB 661 AGATCGGCCCGGAGACCCCTACAAACACCCCGCTGTTCCGCCATCAAGAAGAAGACAGCA 720
 QY 878 CCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGG 937
 DB 721 CCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGG 780
 QY 938 AGTGCAGCTGGGATCCCGGCGGCGGCGCTGAAGAAGAAGAAGAGCGTGAACGCTGC 997
 DB 781 AGTGCAGCTGGGATCCCGGCGGCGGCGCTGAAGAAGAAGAAGAGCGTGAACGCTGC 840
 QY 998 TGGACGTGGCGCGCTACTTTCAGCGTGGCGGCGGCGCTTCCGCAAGTACACCG 1057
 DB 841 TGGACGTGGCGCGCTACTTTCAGCGTGGCGGCGGCGCTTCCGCAAGTACACCG 900
 QY 1058 CTTTACCATCCCGAGATCAACAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1117
 DB 901 CTTTACCATCCCGAGATCAACAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
 QY 1118 TGCCCCAGGCTGGAAGGGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1177
 DB 961 TGCCCCAGGCTGGAAGGGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 QY 1178 AGCCCTTCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1237
 DB 1021 AGCCCTTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1074
 QY 1238 TGGCAGCGAGCTGGAGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1297
 DB 1075 TGGCAGCGAGCTGGAGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1134
 QY 1298 TGCTGGCGTGGGCTTACACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1357
 DB 1135 TGCTGGCGTGGGCTTACACACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1194
 QY 1358 GGATGGGCTACGAGCTGCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1417
 DB 1195 GGATGGGCTACGAGCTGCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1254
 QY 1418 AGGAGCTGGACCGTGAACGACATCCAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1477
 DB 1255 AGGAGCTGGACCGTGAACGACATCCAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1314
 QY 1478 AGATCTACCGCGGATCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1537
 DB 1315 AGATCTACCGCGGATCAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1374
 QY 1538 TGACCGGATCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1597
 DB 1375 TGACCGGATCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1434
 QY 1598 TCCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1657
 DB 1435 TCCTGAAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1494
 QY 1658 TCCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1717
 DB 1495 TCCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1554
 QY 1718 TGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1777
 DB 1555 TGAAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1614
 QY 1778 CCGAGCGG 1837
 DB 1615 CCGAGCGG 1674
 QY 1838 TCCGCGCTGCCCATCCAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1897
 DB 1675 TCAAGCTGCCCATCCAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1734

Qy	1998	CCTGGATCCCGAGTCGGGAGTTCTGTGAACACCCCCCCCTGGTGAAGCTGTGTGATCCAGC	1957
Db	1735	CCTGGATCCCGAGTCGGGAGTTCTGTGAACACCCCCCCCTGGTGAAGCTGTGTGATCCAGC	1794
Qy	1958	TGAGAAGGAGCCCATCATCTGCGCGCGAGACCTTCTACGTGGACGGCGCGCCACACCGCG	2017
Db	1795	TGAGAAGGAGCCCATCTGTGGCGCGAGACCTTCTACGTGGACGGCGCGCCACACCGCG	1854
Qy	2018	AGACCAAGATTCGCAAGCGCGCTACGTGACCGACCGGGGCGGCAGAAAGATCGTGAGCC	2077
Db	1855	AGACCAAGCTGGCAAGCGCGCTACGTGACCGACCGGGGCGGCAGAAAGTGGTGAGCA	1914
Qy	2078	TGACCGGAGACCAACACCAAGACAGACGAGCTGCAGGCCATCCAGCTGGCCCTGCAGSACA	2137
Db	1915	TCGCCGACACCAACCAACCAAGACGAGCTGCAGGCCATCCACCTGGCCCTGCAGSACA	1974
Qy	2138	GCGCACCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCATCTCAGGCC	2197
Db	1975	CGGCCTTGAGGTGAACATCTGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCC	2034
Qy	2198	AGCCGCACAAGAGCGAGCGAGCTGTGTGAACACAGATCATCGACGAGCTGATCAAGAAGG	2257
Db	2035	AGCCGCACAAGAGCGAGCGAGCTGTGTGAGCCAGATCATCGACGAGCTGATCAAGAGG	2094
Qy	2258	AGAAGGTGTACTGTAGCTGGGTGCCCCGCCAACAGGGCATCGCGCGCAACGACGATCG	2317
Db	2095	AGAAGGTGTACTTGGCTGGGTGGCCGCCCAAGAGGCATCGGGGCAACGACGAGTGG	2154
Qy	2318	ACAAGCTGTGAGCANGGCATCCGACAGGTGCTGTCTTGGACGGCATCGATGGCGGCA	2377
Db	2155	ACAAGCTGTGTAGCGCGCGCATCCGCAAGGTGCTGTCTTGAACGGCATCGATGGCGGCA	2214
Qy	2378	TCGTGATCTACCACTACATGGACGACCTGTACGTGGGCAGCGGGCGCCCTAGGATCGATT	2437
Db	2215	TCGTGATCTACCAGTACATGAGACCTGTACGTGGGCAGCGGGCGCCCTAGGATCGATT	2274
Qy	2438	AAAAGCTTCCCGGGGTAGCACCGGTGAATTTC	2469
Db	2275	AAAAGCTTCCCGGGGTAGCACCGGTGAATTTC	2306

RESULT 12

ABK91624

AX
AC

DT 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #176.

HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte CTL.

OS Synthetic.

AX WO200232943-A2.

25-APR-2002.

AA
PF 14 - AUG - 2001; 2001WO-US25721.

XX
PR 14-AUG-2000: 2000US-225097P

PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P

XX
PA (JCSH) US DEPT HEALTH & HUM

PA (CHAD/) CHADRABARTI B K.
YY

PI Nabel GJ, Huang Y;
yy

DR WPI; 2002-452382/48

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure; Page 791-794; 794pp: English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

Sequence 9785 BP; 2360 A; 2843 C; 2731 G; 1851 T; 0 other;

Query Match	81.8%;	Score 2020.4;	DB 24;	Length 9785;
Best Local Similarity	91.6%;	Pred. No. 8.4e-245;		

Qy	14	TGGCCGAGGCATGAGCCAGGCCACACAGCCGCGAATCTCTGATGAGCGCCAGCAACTTCA	73
Db	2967	TGGCCGAGGCATGAGCCAGGCCACACAGCCGCGAATCTCTGATGAGCGCCAGCAACTTCA	3026
Qy	74	AGGCCCCAAAGCGCATCATCAAGTGTTCAACTGCGGCGAAGGAGGCGCCACATCGCCCGCA	133
Db	3027	AGGCCAGAAAGCG--CATCAAGTGTTCAACTGCGGCGAAGGAGGCGCCACATCGCCCGCA	3088
Qy	134	ACTCGCGCGCCCGCGCAAGAGAGGCTGCTGGAAGTTCGGGCAAGGAGGCGCCACCAAGATGA	193
Db	3084	ACTCGCGCGCCCTGCGCAAGAGAGGCTGCTGGAAGTTCGGGCAAGGAGGCGCCACCAAGATGA	314
Qy	194	AGGACTCACCGAGCGCGAGGCCAACTCTTCCGCGAGGACCTGGCCTTCCGCCCAAGGCA	253
Db	3144	AGGACTCACCGAGCGCGAGGCCAACTCTTCCGCGAGGACCTGGCCTTCCGCCCAAGGCA	3197
Qy	254	AGGCCCGGAGTTCCCGAGCGAGCGAGAACCGCGGCCAACAGCCCCACACAGCGCGAGCTGC	313
Db	3198	AGGCCCGGAGTTCTTTCAGAGCAGACAGAGGCCAACAGCCCCACACAGCGCGAGCTGC	3255
Qy	314	AGGTGCGCGG----CGACAACCCCGAGAGCGAGCGCGCGCGCGAGCGCCAGGCGCA---	364
Db	3258	AGGTTTGGGGAAGAGACAACAATCCCTCTCAGAGCAGAGCGCGGATAGCAAGGAATG	3317
Qy	365	---CCCTGAACTTCCCGCCAGATCACCTCTGGCGAGCGCCCTCGTGTGAGCATCAAGTGG	421
Db	3318	TATCTTTAGCTTCCCTCAGATCACCTTTTGGCAGCGACCCCTCGTCACATAAAGATAG	3377
Qy	422	CGGCGCAGATCAAGGAGGCGCTCTCGACACCGCGCGCGCGACACCGTGTGTGGAGGAGA	481
Db	3378	GGGGCCAGCTGAAGAGGCGCCTTCTAGACACCGCGCGCGCGACACCGTGTGTGGAGGAGA	3437
Qy	482	TGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGCATCGCGGCTTCATCAAGG	541
Db	3438	TGAACCTGCCCGGCGCTGGAAGSCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGG	3497
Qy	542	TGGCGCATGACACAGATCTGATCGAGATCTGGGCAAGAGGCCATCGGCGACCGTGC	601
Db	3498	TGGGCCAATGACACAGATCTGATCGAGATCTGGGGCCACAAAGGCCATCGGCGACCGTGC	3557
Qy	602	TGATGCGCCCCACCCCGTGAACATCATCGGCGCGAAACTGCTGACCCAGCTGGGCTCSA	661
Db	3558	TGGTGGCCCCACCCCGTGAACATCATCGGCGCGAACTGCTGACCCAGATCGGCTGCA	3617
Qy	662	CCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGCATGG	721
Db	3618	CCCTGAACTTCCCCATCAGCCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGCATGG	3677

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or human
PT immunisation for acquired immunodeficiency syndrome or protein
PT immunodeficiency virus infection -
XX
PS Disclosure: Page 739-741; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9189 BP; 2232 A; 2671 C; 2524 G; 1762 T; 0 other;

Query Match 81.8%; Score 2018.6; DB 24; Length 9189;
Best Local Similarity 91.4%; Pred. No. 1.4e-244;
Matches 2190; Conservative 0; Mismatches 184; Indels 21; Gaps 4;

QY 14 TGCGCGAGGCGCATGAGCCAGG---CCACCAGCGCCCAACATCCTGTATGACGCGCAGCAACT 70
DB TGCGCGAGGCGCATGAGCCAGGTCACCAAGCGCCCAACATCCTGTATGACGCGCGCAACT 3026
QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
DB TCCGCAACCAAGCGCATCGTGAAGTGTCTTCAACTGCGGCAAGGAGGGCCACATCGCCC 3086
QY 131 GCAACTGCGCGCGCCCGCGCAAGAGGGTGTGTTGAAGTGTGCGCAAGGAGGGCCACACAGA 190
DB GCAACTGCGCGCGCCCGCGCAAGAGGGTGTGTTGAAGTGTGCGCAAGGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGACCGAGCGCGCCAGCAACTTCTTCCGCGAGGACCTGGCTTCCCGCCAGG 250
DB TGAAGGACTGACCGAGCGCGCCAGCAACTTCTTCCGCGAGGACCTGGCTTCCCGCCAGG 3200
QY 251 GCAAGGCCCGCGAGTTCGCCAGCGAGCAGAACCGCGCCCAACAGCCGCCACACGCGCGAGC 310
DB GCAAGGCCCGCGAGTTCGCCAGCGAGCAGAACCGCGCCCAACAGCCGCCACACGCGCGAGC 3260
QY 311 TGCAGTGTGCGCG- - - - -CGACACCCCGCGAGCGAGCGCGCGCGCGCGCGCGCGCGCA 364
DB TGCAGTGTGCGCGAGAGAGACACAACTCTCTCAGAAAGCAGGAGCGCGATAGACAAAGGAA 3320
QY 365 - - - - -CCCTGAACTTCCCGCAGATCACCTGTGCGAGCGCGCCCTGGTGAGCATCAAGG 418
DB CTGTATCTTTCCTTCCCTCAGATCCTCTTTGGCAGGACCCCTTCGTCACAAATAGAAG 3380
QY 419 TGGCGCGCGCAGATCAAGAGGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCGG 478
DB TAGGGGGCCAGCTGAAGAGGGCCCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGG 3440
QY 479 AGATGAGCCTGCCCGCAGAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCGCTTCATCA 538
DB AGATGAGCCTGCCCGG 3500
QY 539 AGGTGCGCGCAGTACGACCAAGATCTCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCAACG 598
DB AGGTGCGCGCAGTACGACCAAGATCTCTGATCGAGATCTGCGGCGCAAGAGGCGCATCGGCAACG 3560
QY 599 TGTGTATCGGCGGCT 658
DB TGTGTATCGGCGGCT 3620
QY 659 GCACCTGTAACCTCCCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGCA 718
DB GCACCTGTAACCTCCCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGCA 3680

QY 719 TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
DB TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGAGTGG 3740
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCGCAAGATCAACAAAGATCGGCGCGCGAGAACCCCT 838
DB AGATCTGACCCGAGATGGAGAAGGAGGCGCAAGATCAACAAAGATCGGCGCGCGAGAACCCCT 3800
QY 839 ACAACACCCCGGTGTGGCCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGGTGG 898
DB ACAACACCCCGGTGTGGCCATCAAGAAAGAGGAGACACCAAGTGGCGCAAGCTGGTGG 3860
QY 899 ACTTCGCGAGCTGAACAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCGCC 958
DB ACTTCGCGAGCTGAACAAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCGCC 3920
QY 959 ACCCGCGCGCTGAAGAAGAGAGCTGACCGTGTGGAGTGGGCGGAGCGCTACT 1018
DB ACCCGCGCGCTGAAGAAGAGAGCTGACCGTGTGGAGTGGGCGGAGCGCTACT 3980
QY 1019 TCAGCTGCGCCCTGGAGCGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCGAGCATCA 1078
DB TCAGCTGCGCCCTGGAGCGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCGAGCATCA 4040
QY 1079 ACAACGAGACCCCGCGCATCCGCTACCACTACAACGTGCTGCCCGAGGCTGGAAGGCA 1138
DB ACAACGAGACCCCGCGCATCCGCTACCACTACAACGTGCTGCCCGAGGCTGGAAGGCA 4100
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCCGCGCGCAAC 1198
DB GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCCGCGCGCAAC 4160
QY 1199 CCAGATCTGTGATCTTACCAGTACATGGAGCAGCTGTACGTGGGCGAGCGACCTGGAGTCG 1258
DB CCAGATCTGTGATCTTACCAGTACATGGAGCAGCTGTACGTGGGCGAGCGACCTGGAGTCG 4220
QY 1259 GCGAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCCGTGGGCTTCACCA 1318
DB GCGAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCCGTGGGCTTCACCA 4280
QY 1319 CCGCCGACAGAGCAGCAGGAGCGCGCTTCTCTGTGGATGGGCTAGAGCTGCACC 1378
DB CCGCCGACAGAGCAGCAGGAGCGCGCTTCTCTGTGGATGGGCTAGAGCTGCACC 4340
QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGGAGAGGAGAGCTGGAACG 1438
DB CCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGGAGAGGAGAGCTGGAACG 4400
QY 1439 ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCGGCATCAAG 1498
DB ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCGCAGCCAGATCTACCCGGCATCAAG 4460
QY 1499 TGGCCGAGCTGTCAAGCTGCTCGCGGCGCAGGCGCTGACCGAGCATCGTCCCGCTGA 1558
DB TGGCCGAGCTGTCAAGCTGCTCGCGGCGCAGGCGCTGACCGAGCATCGTCCCGCTGA 4520
QY 1559 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTTCCGCGGAGCGCGTGCAG 1618
DB CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTTCCGCGGAGCGCGTGCAG 4580
QY 1619 GCGTGTACTAGACCCCGCAGCAAGGACTTGGTGGCGCAGATCCAGAGCAGGCGCAGGAC 1678
DB GCGTGTACTAGACCCCGCAGCAAGGACTTGGTGGCGCAGATCCAGAGCAGGCGCAGGAC 4640
QY 1679 AGTGGACCTACAGATCTTACCAGGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCA 1738
DB AGTGGACCTACAGATCTTACCAGGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCA 4700
QY 1739 AGATGCGCACCGCCACACAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAGATCG 1798
DB AGATGCGCACCGCCACACAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAGATCG 4760

Qy 1799 CCATGGAGAGCATCGTGTGATCTGGGCGAAGACCCCAAGTTCGGCTGCCATCCAGAGG 1858
Db 4761 CCACCGAGAGCATCGTGTGATCTGGGCGAAGACCCCAAGTTCAGCTGCCATCCAGAGG 4820
Qy 1859 AGACCTGGGAGACCTGTGTGACCCACTACTGGCAGGCCACCTGTGATCCCGAGTGGGAGT 1918
Db 4821 AGACCTGGGAGGCTGTGTGACCCAGTACTGTGGCAGGCCACCTGTGATCCCGAGTGGGAGT 4880
Qy 1919 TCGTGAACACCCCCCTGTGTGAAGTGTGTACAGCTGGAGAGGAGGCCATCATCG 1978
Db 4881 TCGTGAACACCCCCCTGTGTGAAGTGTGTACAGCTGGAGAGGAGGCCATCATCG 4940
Qy 1979 GCGCCGAGACCTTCTACGTGTGACCGGCGGCCCAACCGCAGACCAAGATCGCGAAGGCCG 2038
Db 4941 GCGCCGAGACCTTCTACGTGTGACCGGCGGCCCAACCGCAGACCAAGATCGCGAAGGCCG 5000
Qy 2039 GCTACGTGACCGACCGCGCGCGGCGGAGAGATCGTGTGACCGTGAACGAGACCAACACAGA 2098
Db 5001 GCTACGTGACCGACCGCGCGCGGCGGAGAGATCGTGTGACCGTGAACGAGACCAACACAGA 5060
Qy 2099 AGACCGAGCTGCAGCGCATCCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCG 2158
Db 5061 AGACCGAGCTGCAGCGCATCCACTGGCCCTGCGAGGACAGCGGCGTGGAGTGAACATCG 5120
Qy 2159 TGACCGACAGCAGTACGCCCTGGGCGATCATCTCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 2218
Db 5121 TGACCGACAGCAGTACGCCCTGGGCGATCATCTCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 5180
Qy 2219 AGCTGGTGAACCATCATCTGAGCAGCTGTATCAAGAGGAGGAAGTGTACCTGAGCTGGG 2278
Db 5181 AGCTGGTGAACCATCATCTGAGCAGCTGTATCAAGAGGAGGAAGTGTACCTGAGCTGGG 5240
Qy 2279 TGCCCGGCCCAAGGGCATCGCGGCGCAACGAGCAGATCGCAAGCTGTGTGAGCAAGGSCA 2338
Db 5241 TGCCCGGCCCAAGGGCATCGCGGCGCAACGAGCAGATGTGAGCGGCGTGTGTGAGCGCGGSCA 5300
Qy 2339 TCCGCAAGGTGCTTCTCTGAGCGGCATCGATGGCGGCATCGTATGATCTACCAAGTA 2393
Db 5301 TCCGCAAGGTGCTTCTCTGAGCGGCATCGATGGCGGCATCGAAGGCCAGGAGGACCGAGAA 5355

RESULT 14

ABK91619
ID ABK91619 standard; DNA; 9167 BP.
XX AC ABK91619;
XX DI 14-AUG-2002 (first entry)
XX DE Modified HIV protein-encoding plasmid DNA #171.
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; Aids;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX OS Synthetic.
XX PN WO200232943-A2.
XX PD 25-APR-2002.
XX PF 14-AUG-2001; 2001WO-US25721.
XX PR 14-AUG-2000; 2000US-225097P.
XX PR 14-NOV-2000; 2000US-252115P.
XX PR 28-MAR-2001; 2001US-279257P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PA (CHAD/) CHADABARTI B K.
XX PI Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
DR

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX PS Disclosure; Page 775-778; 794pp; English.
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX SQ Sequence 9167 BP; 2225 A; 2668 C; 2517 G; 1757 T; 0 other;

Query Match 81.7%; Score 2017; DB 24; Length 9167;
Best Local Similarity 91.4%; Pred. No. 2.3e-244;
Matches 2189; Conservative 0; Mismatches 185; Indels 21; Gaps 4;
Qy 14 TGGCGGAGCCCATGAGCCAGG---CCACCAGCGCCAAACATCTGATCGAGCGCACT 70
Db 2964 TGGCGGAGCCCATGAGCCAGGTCACCAACAGCGCCCACTATGATGAGCGGCACT 3023
Qy 71 TCAAGGGGCCCAAGCGCATCATCAAGTCTTCAACTGGCGGAAGGAGGCCACATCTGCC 130
Db 3024 TCCGCAACCGCGCNAAGATCGTGAAGTCTTCACTGCGGCAAGGAGGCCACACGCC 3083
Qy 131 GCACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGGCAAGAGGCGCACACAGA 190
Db 3084 GCACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGGCAAGAGGCGCACACAGA 3143
Qy 191 TGAAGGACTGACCGAGCGCGAGGCCAACTTCTCCGCGAGGACCTGGCTTCCCGCAGG 250
Db 3144 TGAAGGACTGACCGAGCGCGAGGCCAACTTCTCCGCGAGGACCTGGCTTCCCGCAGG 3197
Qy 251 GCAAGGCGCGGAGTTCCCGCAGCAGAGAACCGCGCCCAACAGCCCGCCCGCGAGC 310
Db 3198 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACAGCCACAGCCCGCCAGAGAGC 3257
Qy 311 TGCAGGTGCGCGG-----CGACAACCGCGCAGAGCGCGCGCGCGCGAGCGCGCA 364
Db 3258 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAAGCAGAGCGCGCGATAGCAAGAA 3317
Qy 365 -----CCCTGAACCTTCCCGCAGATCACCTCTGGCAGCGCCCGCTGTGAGCATCAAG 418
Db 3318 CTGTATCTTCTAGCTTCCCTCAGATCATCTTTGGCAGCGACCCCTCTGCAATAAAGA 3377
Qy 419 TGGCGGCGCAGATCAAGGAGCGCTTCTGGACACCGCGCGCGCGCGCACACCGTGTGGAG 478
Db 3378 TAGGGGCGCAGCTGAAGGAGCGCTTCTAGACACCGCGCGCGCGCACACCGTGTGGAG 3437
Qy 479 AGATGAGCTGCGCGGCAAGTGGAGCCCAAGATGATCGGGGCGATCGGGCGCTTCATCA 538
Db 3438 AGATGAACCTGCGCGCGCTGGAAGCCCAAGATGATCGGGGCGATCGGGCGCTTCATCA 3497
Qy 539 AGGTGCGCGCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 3498 AGGTGCGCGCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3557
Qy 599 TGCTGATCGCGCGCCACCGCTGAAACATCATCGGCGCGCAACATGCTGACCCAGCTGGCT 658
Db 3558 TGCTGCTGGCGCCCGCCCGCTGACATCATCGGCGCGCAACCTGCTGACCCAGATCGCT 3617
Qy 659 GCACCTGAACCTTCCCATCATCGCCCGCCATCGAGACCGCTGCCCTGAAGCTGAAGCCCGCA 718
Db 3618 GCACCTGAACCTTCCCATCATCGCCCGCCATCGAGACCGCTGCCCTGAAGCTGAAGCCCGCA 3677

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX
PS Disclosure; Page 760-763; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9170 BP; 2225 A; 2669 C; 2519 G; 1757 T; 0 other;

Query Match 81.7%; Score 2017; DB 24; Length 9170;
Best Local Similarity 91.4%; Pred. No. 2.3e-244;
Matches 2189; Conservative 0; Mismatches 185; Indels 21; Gaps 4;

QY 14 TGGCCGAGGGCCATGAGCCAGG---CCACCAGCGCCAACTCCTGATGCGAGCGCAGCAACT 70
DB TGGCCGAGGGCCATGAGCCAGGTTGACCAACAGCGCCACCATCATGATGAGCGCGGCAACT 3026

QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGCCTTCAACTGCGGAAGAGGGCCACATCGCC 130
DB TCCCAACACAGCGCAAGATGCTGAAGTGCCTTCAACTGCGGAAGAGGGCCACACCGGCC 3086

QY 131 GCAACTGCGCGCCCGCCGCAAGAGGGCTCTGGAGTCCGGCAAGAGGGCCACACAGA 190
DB GCAACTGCGCGCCCGCCGCAAGAGGGCTCTGGAGTCCGGCAAGAGGGCCACACAGA 3146

QY 191 TGAAGGACTGCACGAGCGCCAGCGCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGG 250
DB TGAAGGACTGCACGAGCGCCAGCGCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGG 3200

QY 251 GCAAGGCGCGGATTCGCCAGCGAGGAGCAACCGCGCAACAGCCACAGCGCGCGAGC 310
DB GGAAGGCGCGGATTTCTTTCAGAGCAGACCAAGAGCCACAGCCCGCCACAGAGAGC 3260

QY 311 TGCAGGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCCAGGGCA 364
DB TTCAGGTTGGGGAAGAGACAACAACCTCCCTCTCAGAGCAGGAGCGCGATAGACAAGGAA 3320

QY 365 -----CCCTGAATTCCTCCAGATCACCTGTGGCAGCGCCCTGTGAGCATCAAGG 418
DB CTGTATCTTCTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCAATTAAGA 3380

QY 419 TGGCGCGCGAGATCAAGAGCGCCCTGTGGACACCGCGCGCGAGCACCGGTGTGGAGG 478
DB TAGGGCGCGAGCTGAAGAGGCGCCCTTTAGACACCGCGCGCGAGCACCGGTGTGGAGG 3440

QY 479 AGATGAGCTTCCCGCAAGTGGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB AGATGAGCTTCCCGCGCGCTGGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 3500

QY 539 AGGTGCGCCAGTAGACACAGATCTGTATCGAGATCTCGCGCAAGAGCCATCGCGCACCG 598
DB AGGTGCGCCAGTAGACACAGATCTGTATCGAGATCTCGCGCCACAAAGGCCATCGCGCACCG 3560

QY 599 TGTGATCGGCGCCACCCCGTGAACATCATPCGCGCCGCAACATGCTGACCCAGCTGGGCT 658
DB TGTGATCGGCGCCACCCCGTGAACATCATPCGCGCGCAACATGCTGCTGACCCAGATCGGCT 3620

QY 659 GCACCTGGAATTCCTCCATCAGCCCCATCAGAGCCGTGCGCGGTGAAGCTGAGCGCCGCA 718
DB GCACCTGGAATTCCTCCATCAGCCCCATCAGAGCCGTGCGCGGTGAAGCTGAGCGCCGCA 3680

QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCC 778
DB TGGACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGTTGG 3740

QY 779 CCATCTCGAGGAGATGGAGAAAGGAGGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 838
DB AGATCTGCACCGAGATGGAGAAAGGAGGCAAGATCAGCAAGATCGGCCCGCGAGAACCCCT 3800

QY 839 ACAACACCCCGCTTCCCATCAAGAAAGGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
DB ACAACACCCCGCTTCCCATCAAGAAAGGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 3860

QY 899 ACTTCCGCGAGCTGAACAAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGCATCCCCC 958
DB ACTTCCGCGAGCTGAACAAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGCATCCCCC 3920

QY 959 ACCCGCGCGGCTTGAAGAAAGAGAGAGCGTGCAGCTGCTGGAGTGGCGGACGCTACT 1018
DB ACCCGCGCGGCTTGAAGCAGAGAGAGAGCGTGCAGCTGCTGGAGTGGCGGACGCTACT 3980

QY 1019 TCAGCGTCCCGCTGGACGAGGACTTCCCAAGTACACCGCCCTTCCATCCCCAGCATCA 1078
DB TCAGCGTCCCGCTGGACAAGGACTTCCCAAGTACACCGCCCTTCCATCCCCAGCATCA 4040

QY 1079 ACAACGAGACCCCGGCTATCCGCTACCAAGTACAAAGTGTGCTGCCCGAGGCTGGAGGGCA 1138
DB ACAACGAGACCCCGGCTATCCGCTACCAAGTACAAAGTGTGCTGCCCGAGGCTGGAGGGCA 4100

QY 1139 GCGCCAGCATCTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCCCGCAGACC 1198
DB GCGCCAGCATCTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCAGCAGAAC 4160

QY 1199 CCGAGATCGTATCTACCAGTACATGAGCAGCTGTAGCTGGGAGCGACCTGGAGATCG 1258
DB CCGAGATCGTATCTACCAGTACATGAGCAGCTGTAGCTGGGAGCGACCTGGAGATCG 4220

QY 1259 GCCAGCAGCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGGCTTCACCA 1318
DB GCCAGCAGCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGGCTTCACCA 4280

QY 1319 CCGCCGACAAAGACACACAGAGAGGAGCGCCCTTCTGTGTGATGGGTACGAGTGCACC 1378
DB CCGCCGACAAAGACACACAGAGAGGAGCGCCCTTCTGTGTGATGGGTACGAGTGCACC 4340

QY 1379 CCGACAAGTGCACGTCGACGCCCATCGAGCTGCCGCGAAGAGAGCTGGACCGTGAACG 1438
DB CCGACAAGTGCACGTCGACGCCCATCGAGCTGCCGCGAAGAGAGCTGGACCGTGAACG 4400

QY 1439 ACATCCAGAAAGCTGGTGGGCAAGCTGAACTGGGCGAGCGAGATCTACCCCGCATCAAGG 1498
DB ACATCCAGAAAGCTGGTGGGCAAGCTGAACTGGGCGAGCGAGATCTACCCCGCATCAAGG 4460

QY 1499 TGGCGCAGCTGTCAAGCTGTGCGCGCGCCAGGCGCTGACCGACATCGTGGCCCTGA 1558
DB TGGCGCAGCTGTCAAGCTGTGCGCGCGCCAGGCGCTGACCGACATCGTGGCCCTGA 4520

QY 1559 CCGAGGAGCGCGAGCTGGAGCTGCCGAGAACCGCGAGATCTGCGCGAGCGCGTGCAGG 1618
DB CCGAGGAGCGCGAGCTGGAGCTGCCGAGAACCGCGAGATCTGCGCGAGCGCGTGCAGG 4580

QY 1619 GCGTGTACTAGACCCCGAGCAAGGACCTGGTGGCGCGAGATCCAGAGCAGGCGCCACGACC 1678
DB GCGTGTACTAGACCCCGAGCAAGGACCTGGTGGCGCGAGATCCAGAGCAGGCGCCACGACC 4640

QY 1679 AGTGGACCTACAGATCTACCAAGGAGCGCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 1738
DB AGTGGACCTACAGATCTACCAAGGAGCGCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 4700

QY 1739 AGATGCGCACCGCCCAACCAACAGCTGAGAGAGCTGACCGAGCGCGTGCAGAGATCG 1798
DB GCATGAAGGGCGCCCAACCAACAGCTGAGAGAGCTGACCGAGCGCGTGCAGAGATCG 4760

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 11:44:48 : Search time 3394.25 Seconds
(without alignments)
11780.710 Million cell updates/sec

Title: US-09-610-313-30

Perfect score: 2469

Sequence: 1 gtcgacgcacccatggccga.....gggctagcaccgggaattc 2469

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpi:**

7: em_estro:**

8: em_htc:**

9: gb_estl:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_other:**

26: em_gss_pro:**

27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.8	3.7	2598	11 AY103647	AY103647 Zea mays
2	85.6	3.5	951	13 BM321451	BM321451 rockefell
3	80.2	3.2	1132	13 BM320864	BM320864 rockefell
4	79.6	3.2	1165	13 BM320900	BM320900 rockefell
5	79	3.2	867	13 BM321430	BM321430 rockefell
6	76.6	3.1	1550	13 BM321022	BM321022 rockefell

7	75.4	3.1	2299	11 AY106831	AY106831 Zea mays
8	75	3.0	862	13 BM321023	BM321023 rockefell
9	74.4	3.0	3134	11 AY109500	AY109500 Zea mays
10	73.2	3.0	853	13 BM321393	BM321393 rockefell
C 11	72.8	2.9	566	13 BM587428	BM587428 170006873
C 12	72.4	2.9	801	13 BJ375394	BJ375394 BJ375394
13	71.2	2.9	646	13 BM645449	BM645449 170006873
14	71.2	2.9	724	13 BM584191	BM584191 170006872
C 15	71	2.9	925	17 CNS0091P	AL053013 Drosophila
16	70.6	2.9	861	10 BEG36696	BEG36696 rockefell
17	70	2.8	494	12 BG349122	BG349122 947029804
18	69.6	2.8	546	14 BQ167279	BQ167279 WHE0062_D
19	69.4	2.8	640	10 BE601575	BE601575 HVSMEH009
20	69.4	2.8	926	14 BQ949605	BQ949605 AGENCOURT
21	69.2	2.8	693	12 BG850115	BG850115 1024028A1
22	69	2.8	901	9 AL551163	AL551163 AL551163
23	68.8	2.8	813	13 BI831786	BI831786 603079057
24	68.8	2.8	819	10 BE040830	BE040830 OF12C07 O
25	68.4	2.8	937	13 BM321010	BM321010 rockefell
26	68.4	2.8	1032	10 BE040732	BE040732 OF05G08 O
27	68.4	2.8	1101	17 CNS00397	AL063912 Drosophila
28	68.4	2.8	1333	13 BM320976	BM320976 rockefell
29	68.2	2.8	596	10 AW909955	AW909955 ur78b09 Y
30	68.2	2.8	799	13 BJ361740	BJ361740 BJ361740
31	68	2.8	500	13 BM372120	BM372120 EBR003_SQ
32	68	2.8	530	14 BQ460810	BQ460810 HA05N15r
33	68	2.8	536	14 BQ469029	BQ469029 HM03C13r
34	68	2.8	538	13 BM368580	BM368580 EBR008_SQ
35	68	2.8	540	9 AJ471121	AJ471121 AJ471121
36	68	2.8	550	10 AV912864	AV912864 AV912864
37	68	2.8	566	14 BQ464692	BQ464692 HF02P20r
38	68	2.8	578	13 BM377112	BM377112 EBR005_SQ
39	68	2.8	579	12 BF253640	BF253640 HVSMEF000
40	68	2.8	582	13 BM372871	BM372871 EBR004_SQ
41	68	2.8	588	14 BQ765498	BQ765498 EBR003_SQ
42	68	2.8	600	9 AL508778	AL508778 AL508778
43	68	2.8	625	14 BQ768295	BQ768295 EBR008_SQ
44	68	2.8	641	12 BG908906	BG908906 Tatr1172B
45	68	2.8	656	13 BM370222	BM370222 EBR008_SQ

ALIGNMENTS

RESULT 1
AY103647
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

AY103647
Zea mays
PCO142084 mRNA sequence.
AY103647
AY103647.1 GI:21206725
HTC.
Zea mays.
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 2598)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 2598)
Coe,E.C.
Direct Submission
Submitted (25-APR-2002) Maize Mapping Project, University of
Missouri, Columbia, MO 65211, USA
Location/Qualifiers
1..2598
/organism="Zea mays"
/db_xref="MaizeDB:638378"
/db_xref="taxon:4577"
/clone="PCO142084"

linear HTC 25-MAY-2002

ACCESSION BM321451
 VERSION BM321451.1 GI:18055857
 KEYWORDS EST.
 SOURCE Mastigamoeba balamuthi.
 ORGANISM Mastigamoeba balamuthi.

REFERENCE Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

AUTHORS 1 (bases 1 to 951)
 Baptiste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W.,
 Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and
 Philippe, H.

TITLE The analysis of 100 genes supports the grouping of three highly
 divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

MEDLINE 21819461

COMMENT Contact: Muller Miklos

Location/Qualifiers

1. 951

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

BASE COUNT 186 a 321 c 303 g 139 t 2 others

ORIGIN

Query Match 3.5%; Score 85.6; DB 13; Length 951;

Best Local Similarity 46.0%; Pred. No. 9.9e-05;

Matches 323; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

1;

Qy 1734 GCGCAAGATGCGACCGCCACACACAGCAGTGAAGCAGTGACCGAGGCGGTGCAGAA 1793

Db 219 GCGCAAGATGCGACCGCCACACACAGCAGTGAAGCAGTGACCGAGGCGGTGCAGAA 278

Qy 1794 GATCGCCATGGAGCATCTGTATCTGGGCAAGACCCCAAGTTCCGCTGCCCATCA 1853

Db 279 CATCGCGAGCGCGGCTCTCGCTTCTCGCTGGAGGCGCAGACCTCCAGGAGTATG 338

Qy 1854 GAAGAGACCTGGAGACCTGGTGACCGACTACTGGCAGGCCACCTGGATCCCGCAGTG 1913

Db 339 GGAGTGCACCTGAAGGCGCTGTGCTCGGCCCTTACCAGGCGCTCAGATCATGTCGA 398

Qy 1914 GGAGTTCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCAT 1973

Db 399 CGACGGCGGTGACCGGACTCTGTATGATCCACAGGGGTTCGCGCGGAGGACAAACCCAA 458

Qy 1974 CATCGCGCCGAGACCTTCTACGTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAA 2033

Db 459 GCTGCTGGAGCAGCAGGAGGCGCTCGAGGAGTTCGCTTCAACACGTCTCAAGCA 518

Qy 2034 GCGCGCTACTGTACGACCGCGCGCGCAGAGTGTGACCTGTGACCGGAGACCAACAA 2093

Db 519 GGTCCAGAGGAGCAGCGCGCTTCTGSCACAAGATCTCTCCCGAGATCCCGGCTGTGAG 578

Qy 2094 CCAGAAGACCGAGCTGCGAGGCAATCCAGCTGGCCCTTCAGGACAGCGCAGGAGTGAA 2153

Db 579 CGAGGAGACGACTGGCGTGTATGAGCTGTACACAGTGTACCGCGGAGGAGGAGTCT 638

Qy 2154 CATGCTGACCGACAGGAGTACGCGCTGGGCATCATCCAGCGCCAGCGCCGACAGAGGGA 2213

Db 639 GTTCCCGCGCTCAACGTCAACGAC---TCTNTCAACAGAGCAGTTCGACACATCTA 695

Qy 2214 GAGGAGCTGTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAG 2273

Db 696 CGGCTGCGGCACATCGCTCATCGACCGGATCAAGCGCGGACCGACGCTGTGCTCGGCG 755

Qy 2274 CTGGGTGCGGCCACAGGCGCATCGCGGCAACGAGCAGATTCACACAGCTGTGTGAGCAA 2333

Db 756 CAAGGTGCGCGCTGCTCGCGGCTACGGCGACGTGGGCAAGGGCTGGCGGCTGCGTGGC 815

Qy 2334 GGGCATCCCAAGGTCTCTTCTGGACGGCATCGATGCGGCGCATGCTATCAGTA 2393

Db 816 CGGCCAGGCTGCGCGTCTCATGTCAGCGAGATCGACCCCATCTGGCGCTCGAGCGCTC 875

Qy 2394 CATGACGACCTGTAGTGGCAGCGCGGCGCTTAGGATCGA 2435

Db 876 GATGCGCGCTTCGAGGTCAACACGCTCGAGGGGGGCTCGA 917

RESULT 3

BM320864

LOCUS

DEFINITION

rockefeller.0.46 Mastigamoeba balamuthi lambda ZAP II Library

Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA

sequence.

BM320864

ACCESSION

VERSION

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

1. 1132

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

BASE COUNT 220 a 413 c 324 g 171 t 4 others

ORIGIN

Query Match 3.2%; Score 80.2; DB 13; Length 1132;

Best Local Similarity 45.3%; Pred. No. 0.00079;

Matches 375; Conservative 0; Mismatches 444; Indels 9; Gaps 2;

Qy 139 CGCGCCCCCGAAGAAGGGCTGTGGAGTGGCGCAAGAGGGCGCCACAGATGAAGGAC 198

Db 79 CGCGCGAGGCGAAGACGAGCTACCGCGCGGCCANCTGGTATCCAGGACAAAGAAC 138

Qy 199 TGCACGAGCGCCAGCGCCAACTTCTTCGCGAGGACCTTGGCTTCCCGCAGGCGCC 258

Db 139 TACACACCCCAAGTACCGCTTCTGTC-----GTCCGCTTCCACCAAGGAGCATCGTC 192

Qy 259 CGCGAGTTCCCGCAGGAGCAGAACCGCGCCCAACAGCCCCACCAGCGCGGAGCTGCAGGTG 318

Db 193 TGCCAGATCGCTTACGCCAAGATCGACGGCGACACATCTCTCGCGCGCCCTACTCGCNC 252

Qy 319 CGCGGCGACACCCCGCAGGAGCGCCGCGCGCAGCGCCAGGCGACCTTGAACCTTCCCC 378

Db 253 GAGTCAACCCGCTCGGCGTCAAGCTCGCGCTGACCAACTACGCGCGCCCTACGCGACT 312

Qy 379 CAGATCACCTGTGGCAGCGCCCTGTGTGAGCATCAAGTGGGCGGCGCAGATCAAGGAG 438

Db 313 GGCCTGTGCTGGCGCGCCGCTGTGCTGAAGAAGCTCAACCTCGACTCCAAAGTACAGGCT 372

Qy 439 GCGCTGTGGACACCGCGCGCGGAGCAGACACCGTGTCTGGAGGAGATGAGCTGCGCGGCAAG 498

```

Db 373 GTCAAGAGTCAACGGCGAGGACTACAGCTCGAGAGCTCGACAGCGGCCCGCGT 432
QY 499 TGAAGCCCAAGATGATCGCGCGCATCGGGGGTTCATCAAGGTGCGGCAGTAGACAGCAG 558
Db 433 TTCAAGGCCCTGCTCGACGTCGGCCTGGTCCGCAACCTCGACTGGCGCCGCGTGTTCGCC 492
QY 559 ATCTGATCGAGATCTGCGCAAGAGGCCATCGGCACCGCTGTGATCGGCCCCACCC 618
Db 493 GCCTCAAGGGCATGTGCGACGGCGCGTCAAGTCCCC--CACAGCGAGACCCGCTTC 549
QY 619 GTGAAGATCATCGCGCGCAACATGCTGACCCAGCTGGGTGCACTTGAATTTCCCATC 678
Db 550 GTCGGCTTCAACGGCGACAGAGAGAGCTCAACGCGCGGTCTTCGCAAGTACATCTTC 609
QY 679 AGCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCGATGACGCGCCCAAGGTGAAG 738
Db 610 GCGCGCACGTCGCGCGCTACATGAAGTCTCTCAAGGAGCAGGACGCGCGCTTCGAC 669
QY 739 CAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTCGAGGAGATGGAG 798
Db 670 CGCCAGTTCTCGCGCTACGCCAAGAGGGTGTACCGCGCGACATGCTCGAAGAGATCTAC 729
QY 799 AAGGAGGCAAGATCAACAGATCGCGCGCGAGAACCCCTTACACACCCCGTGTTCGCC 858
Db 730 ACCGAGGCCACAGCAGATCCGCGCGGACCCGACCTTCGTCCCAAGCGCGCTCGAAG 789
QY 859 ATCAAGAGAGAGACAGCACCAGTGGCGCAAGCTGGTGTGAGTTCGCGAGCTGAACAAG 918
Db 790 CCGAGGGGCCCAAGGCCAAGCACTGGGGCAAGCGAGGTGACGTACCAAGCGCGCAAG 849
QY 919 CGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCCACCGCGCC 966
Db 850 AACCGGTGCCCAGAGAGAGTCCGCTGGGTACCCCGACGCCCC 897

```

```

RESULT 4
BM320900
LOCUS
DEFINITION
  rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
  Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
  sequence.
ACCESSION
  BM320900
VERSION
  BM320900.1 GI:18055306
KEYWORDS
  EST.
SOURCE
  Mastigamoeba balamuthi.
  Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
REFERENCE
  1 (bases 1 to 1165)
  Baptiste, E., Brinkmann, H., Lee, J.A., Moore, D.V., Sensen, C.W.,
  Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and
  Philippe, H.
  The analysis of 100 genes supports the grouping of three highly
  divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  21819461
COMMENT
  Contact: Muller Miklos
  Laboratory of Biochemical Parasitology
  The Rockefeller University
  1230 York Avenue, New York, NY 10021, USA
  Email: mmuller@rockvax.rockefeller.edu
  Insert Length: 1165 Std Error: 0.00
  POLYA=No.
  Location/Qualifiers
    1..1165
      /organism="Mastigamoeba balamuthi"
      /strain="ATCC 30984"
      /db_xref="taxon:108607"
      /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
      /notes="syn: Phreatamoeba balamuthi"
    221 a 426 c 336 g 178 t 4 others

```

FEATURES

```

source
  1..1165
    1..1165
      /organism="Mastigamoeba balamuthi"
      /strain="ATCC 30984"
      /db_xref="taxon:108607"
      /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
      /notes="syn: Phreatamoeba balamuthi"
    221 a 426 c 336 g 178 t 4 others

```

```

BASE COUNT
ORIGIN

```

```

RESULT 5
BM321430
LOCUS
DEFINITION
  rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library

```

```

Query Match 3.2%; Score 79.6; DB 13; Length 1165;
Best Local Similarity 45.1%; Pred. No. 0.00099;
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;

QY 90 CATCAAGTGTTCATCTGCGGCAAGGAGGCGCCACCATGCAAGTGCACCGAGCG 149
Db 46 CGTCAAGAACAAAGCGGTACTTCAAGCGCTTCCAGAGCCAGTTCGTCGCGCGCGAGG 105
QY 150 CAAGAGGGTGTGTAAGTGCAGGAGGAGGCGCCACCATGCAAGTGCACCGAGCG 209
Db 106 CAAGAGGGATACCGCGCGCGCCANCTGGTATCCAGGACAAGAACAGTACACAGCCC 165
QY 210 CCAGGGCAACTTCTCCGCGAGGACCTGGCTTCCCGCAGGGAAGCGCCGAGTTCCC 269
Db 166 CAAGTACCGTTCGTC-----GTCCGCTTACCACAGGAGACATCGTCTGCCAGATCG 219
QY 270 CAGCGAGCAGACGGCGCCCAACAGCCCCACAGCGCGGAGCTGCAAGTGCAGCGGACAA 329
Db 220 CTACGCCAAGATCGAGCGCGACCATCTCTCGCGCGCGCTACTCGACAGAGTCAACCG 279
QY 330 CCCCCGAGCGAGCGCGCGCGCCAGCGCCACCTGAACCTTCCCGCAGATCACCT 389
Db 280 CTTTCGGCGTCAAGTCTGGCTGACCAACTAGCGCGCGCCCTACGCGACTGGCTGTGCT 339
QY 390 GTGCGAGCGCCCTGTGTGAGCATCAAGGTGGCGCGGCGAGTCAAGGAGGCCCTGTGGA 449
Db 340 GCGCGCGCGTGTGTGAAGAGCTCAACCTTACTCCAAGTACGAGGGTGTCAAGAAGGT 399
QY 450 CACCGCGCGCGAGACACCGTGTGAGGAGATGAGCTGCCCGGCAAGTGGAGGCCCAA 509
Db 400 CAACGCGCGAGAGTACAACTCGAGGAGCTCGACGAGCGCGCGCGCTTCAAGGCC-- 457
QY 510 GATCATCGCGCGCATCGCGCTTCATCAAGTGGCGCGCTACGACAGATCTGTATCGA 569
Db 458 --TGTCGACGTGCGCTGTGTCGACCTCGACTGGCGCGCGCTGTTCGCGCGCTCAA 515
QY 570 GATCTGCGGCAAGAAGGCCATCGGCACCTGTGCTGATCGGCCCGCCACCCCGTGAACATCAT 629
Db 516 GGCATGTGCGAGCGCGCGCTCAACGTCCCCACAGCGAGACCGCTTCGTGCGCTTCAA 575
QY 630 CGGCGGCAAGTGTGACCGAGTGGCTGCACCTGAACTTCCCATCAGCCCATCGA 689
Db 576 CGCGCAAGAAGAGAGTCAACGCGCGCTCTCCGCAAGTACATCTTCGCGCGCGCAGT 635
QY 690 GACGTCGCGGTGAAGTGAAGCGCGCATGGAGCGCGCCCAAGGTGAAGAGTGGCCCT 749
Db 636 CGCGGTATCATGAGCTCTCAAGGAGCGAGCGCGCGCTTCGACCGCGCTTC 695
QY 750 GACCGAGGAGATCAAGCGCGCTGACCGCATCTGCGAGGAGATGGAGAGGAGCAA 809
Db 696 GCGCTACGCCCAAGGAGGTGTACCGCGCATCTCGAGAAGATCTACACCGAGGCCA 755
QY 810 GATCACAAGATGGCGCGCGAGAACCCCTACACACCCCGCTGTTCGCCATCAGAGAA 869
Db 756 -----CAAGCAGATCGCGCGCGCGCGCTTCTGTCCTCCCAAGCGCGCTTGAAGCCGA 809
QY 870 GGACAGCACCAAGTGGCGCGAGCTGGTGGACTTCCGCGAGCTGAACAGCGCACCCAGGA 929
Db 810 GGGCGCCCAAGCCNAGACTGGGGCAAGCGAGCTGACGTACCCAGCGCGCAGAACCG 869
QY 930 CTTCTGGGAGTGCAGCTGGGCGATCCCCCGCGCGCTGGAAGAGAGAGAGCGT 989
Db 870 CGTCGCGCAGAAGAAGTTCGCGCTGGGTACCCCGACGCGCCCAAGAGCCAGTAAATTC 929
QY 990 GACGTCGTGGAGTGGGCGAGCGCTACTTTCAGGCTGCCCTGACCGAGG 1039
Db 930 CGGTGCTGTACACCGCCATCTCCGCTGCGCGTTCGCGTGTGCGG 979

```

```

RESULT 5
BM321430
LOCUS
DEFINITION
  rockefeller.0.1153 Mastigamoeba balamuthi lambda ZAP II Library

```


Mastigamoeba balamuthi cDNA similar to ribosomal protein S4, mRNA sequence.

ACCESSION BM321430
VERSION BM321430.1 GI:18055836

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Mastigamoeba balamuthi

REFERENCE 1 (bases 1 to 867)

AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.

The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

21819461

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 867 Std Error: 0.00

POLYA-No.

FEATURES

source

1. .867

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

BASE COUNT 185 a 298 c 254 g 124 t 6 others

ORIGIN

Query Match 3.2%; Score 79; DB 13; Length 867;

Best Local Similarity 45.7%; Pred. No. 0.0012;

Matches 314; Conservative 0; Mismatches 370; Indels 3; Gaps 1;

QY 58 CAGCGCAGCACTTCAAGGGCCCAAGCGCATCATCAAGTCTTCAACTGGCGGCAAGGAG 117

DB 131 CCGCACAAGATGCGGAGTGCCTCCGGTATCATCTGTCGCCACAGAGTTGAAGTAC 190

QY 118 GGCCACATGCGCCGCACTGCGCGCCCGCCGCAAGAGGCTCTGGAAGTGGCGCAAG 177

DB 191 CGCGTGACCGCGCGTGAAGTCACTCGATCGTATGATGAGCGCTGATCAAGATGACGCGC 250

QY 178 GAGCGCCACCATGAGGAGTGAAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 237

DB 251 AAGGTCCGCAACCGACACGCACTTCCCGCGGGCTTCATGGACGTCGTCTCGATCGACAAG 310

QY 238 GCCTTCCCGCAGGCAAGGCGCGGAGTTCGCCAGCGAGCAGACCGCGCAACAGCGCC 297

DB 311 ACCGACGAGCACTTCCGCTGCTCTAGCACACCAAGGCGCTTCAGGCGCACCGCATC 370

QY 298 ACCAGCGCGAGCTGCAAGT---GCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGAG 354

DB 371 AACTCGACCGAGGCAAGTTCAAGCTCGGCAAGTCCGCGCGTGCAGCTCGGCAACAAG 430

QY 355 CGCCAGGCGACCTTGAACCTTCCCGGATCACTTCCCGGATCACTTCCCGGATCACTTCCCGGATCACT 414

DB 431 GGCATCCGTAACCTTGGTGAACCGAGCGCGCACGATCGCTACCCCAACCCCGGATC 490

QY 415 AAGTGGCGCGCAGATCAAGGAGGCGCTGCTGCGACACCGCGCGCGACACCGTGTG 474

DB 491 AAGTCAACGACGCTCAAGATGACATCGCTGCGCGGCAAGATCATCGATTCGTCAAG 550

QY 475 GAGGAGATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATGCGCGGATCGCGGGTTC 534

DB 551 TTCAGATCGGCAACCTTCGTATGATCATCTACCTGCGGAGCGCAACCTTGGCGGTCGCGGTC 610

QY 535 ATCAAGTGGCGCAGTAGGACCATGCTGATCGAGATCTCGGCGCAAGAGCCATCGCG 594

DB 611 ATTGTGCGCGCAGAGACGACGAGGGTCTGTCGATCATCTCCACGTCAGGAGCGCGTC 670

QY 595 ACCGTGTGATCGCGCCACCCCGTGAACATCATTCGGCGCCGAACATCTGACCCAGCTG 654

DB 671 GGCACACAGTTCGGCGCGCGCTGACCAACGCTTCGTGATCGGCAAGGGGACCAAGTCC 730

QY 655 GGCTGCACCCCTGAACATTCGCCCATCAGCCCCATCGAGACCGTGCCTCGTGAAGCTGAAGCCC 714

DB 731 CTCGTACGCTGCCCGCGGCAAGGCGATCAAGAAGTGCATCATCGAGGAGTTCAGCGCG 790

QY 715 GGCATGACGCGCCCAAGGTGAAGCAG 741

DB 791 CGCCACGCGCACAGGACCGAGCAGGAG 817

RESULT 6

LOCUS BM321022 1550 bp mRNA linear EST 03-JAN-2002

DEFINITION rockefeller.0.1192 Mastigamoeba balamuthi lambda ZAP II Library

Mastigamoeba balamuthi cDNA similar to adenosylhomocysteinase (EC

3.3.1.1), mRNA sequence.

ACCESSION BM321022

VERSION BM321022.1 GI:18055428

KEYWORDS EST.

SOURCE Mastigamoeba balamuthi.

ORGANISM Mastigamoeba balamuthi

REFERENCE 1 (bases 1 to 1550)

AUTHORS Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W., Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and Philippe,H.

The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba

Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)

21819461

Contact: Muller Miklos

Laboratory of Biochemical Parasitology

The Rockefeller University

1230 York Avenue, New York, NY 10021, USA

Email: mmuller@rockvax.rockefeller.edu

Insert Length: 1550 Std Error: 0.00

POLYA-No.

FEATURES

source

1. .1550

/organism="Mastigamoeba balamuthi"

/strain="ATCC 30984"

/db_xref="taxon:108607"

/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"

/note="syn: Phreatamoeba balamuthi"

BASE COUNT 282 a 528 c 495 g 226 t 19 others

ORIGIN

Query Match 3.1%; Score 76.6; DB 13; Length 1550;

Best Local Similarity 45.6%; Pred. No. 0.0032;

Matches 308; Conservative 0; Mismatches 364; Indels 3; Gaps 1;

QY 1761 CGAGTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCGCCATGGAGAGCATCGTATCTG 1820

DB 5 CGGACAGGAGCAGGACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 64

QY 1821 GGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGGAC 1880

DB 65 CGCCTGGAAGGCGGAGAACCTCCAGGAGTACTGGAGTGCACCTGGAAGGCCCTGTGTT 124

QY 1881 CGACTACTGGAGCGCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCGCCCTCGT 1940

DB 125 CGGCCCTTACACAGGCGCTCAGATCATGTCGACACGCGGTGACGCGACTCTGTATGAT 184

QY 1941 GAAGCTGTGTACAGCTGGAGAGGAGGCCCATCATCGCGCCGAGACCTTCTACGTGA 2000

DB 185 CCACAAGGGTTCGCGCGGAGGACACCCCAAGCTGCTGGAGGACGACGAGGCGCTCGA 244

QY 2001 CGGCGCGCCCAACCGGAGACCAAGATCGGAAGCGCGGCTACGTGACCGCGGCGCG 2060

DB 245 GGAGTGCCTGCCTCAACAACGTCTCAAGCAGGTCCAGAAGGAGCAGCCCGGCTTCTG 304


```

Db 141 ATCATCATCAAGCTCAAGCAGCTCATCAAGTTCAAGCTCAAGCAGTTCTAGCTCAAGCAG 82
QY 1228 GACCTGTACGTGGG 1241
Db 81 TTCTAGTTCAAGGG 68

RESULT 13
BM645449 646 bp mRNA linear EST 26-FEB-2002
LOCUS 17000687316588 A.Gam.ad.cdna1 Anopheles gambiae cdna clone
DEFINITION 19600449648747 5', mRNA sequence.
ACCESSION BM645449
VERSION BM645449.1 GI:18944960
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE 1 (bases 1 to 646)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004920 row: B column: 09
Seq primer: M13 Reverse.
FEATURES
source
1..646
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449648747"
/clone_lib="A.Gam.ad.cdna1"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site_1: Sall; Site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen.
cdna inserts >500 bp cloned directionally into pSport 1.
Not 1 site is 3'. Clones available through the Malaria
Research and Reference Reagent Resource Center
(www.malaria.mr4.org)."
BASE COUNT 170 a 194 c 210 g 72 t
ORIGIN
Query Match 2.9%; Score 71.2; DB 13; Length 646;
Best Local Similarity 48.1%; Pred. No. 0.023;
Matches 202; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 1956 GCTGGAGAGAGCGCCATCATCGGCGCGGACCTTCTACGTGGACGGCGCGGCAACCG 2015
Db 49 GTTCGAGATCGACGCCAACCGCATCTCTGAGGTGTGCGGCCGAGGACAGGCGCGGCAA 108
QY 2016 CGAGACCAAGATCGCAAGCGCGGTACGTGACCGACCGCGCGCGGCAAGAGATCGTGAG 2075
Db 109 CCGGAGAGATGCTCATCACCACGACCAAGACCGGTGACGCGCGGACGACATCGAGCG 168
QY 2076 CCTGACCGAGACCAACACAGAGACCGAGCTGCGAGCCATCCAGCTGGCCCTGCAGGA 2135
Db 169 CATGATCAAGGATCGGAGCGGTTTCGCGGACGACGACCAAGAGTGAAGGAGCGCGTGA 228
QY 2136 CAGCGGAGCGAGTGACATCTGTGACCGGACGACGACGATACGCCCTGGGCATATCCAGCG 2195
Db 229 GGCCCGCAACGAGCTGAGAGCTACGCGTACAGCTGAAGAACACGCTCAGCTCGAAGGA 288

```

```

QY 2196 CCAGCCGACAAAGAGCGAGAGCGAGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAA 2255
Db 289 CAAGTGGCGCGAGCGTGTCCGACGACGACAAAGGCCAAGATGGAGGAGCGGATCGACGA 348
QY 2256 GGAGAAGTGTACCTGAGCTGGTCCCGCCACAAAGGCGATCGCGCGGACGAGCAGAT 2315
Db 349 GAAGATCAAGTGGCTGGAGACGACGACGACCGGCGGAAGAGTAGTACAAGAGACGAGAA 408
QY 2316 CGACAAGCTGGTGAAGGCGATCCGCAAGGTGCTGTCTCTGGAGCGCATCGATGGCGG 2375
Db 409 GAAGGAGCTGGAGGACATCTGTGACGCCCATCATTTCCCAAGCTGTACGCGAGCAGTGGCGG 468

RESULT 14
BM584191 724 bp mRNA linear EST 22-FEB-2002
LOCUS 17000687279987 A.Gam.ad.cdna1 blood1 Anopheles gambiae cdna clone
DEFINITION 19600449683331 5', mRNA sequence.
ACCESSION BM584191
VERSION BM584191.1 GI:18872658
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE 1 (bases 1 to 724)
AUTHORS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
TITLE Celera Anopheles gambiae EST project
JOURNAL Unpublished (2002)
COMMENT Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004B2V row: B column: 21
Seq primer: M13 Reverse.
FEATURES
source
1..724
/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449683331"
/clone_lib="A.Gam.ad.cdna1 blood1"
/dev_stage="Adult"
/lab_host="DH10b"
/notes="Vector: pSport1; Site_1: Sall; Site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cdna inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)."
BASE COUNT 183 a 218 c 243 g 80 t
ORIGIN
Query Match 2.9%; Score 71.2; DB 13; Length 724;
Best Local Similarity 48.1%; Pred. No. 0.024;
Matches 202; Conservative 0; Mismatches 218; Indels 0; Gaps 0;

QY 1956 GCTGGAGAGAGCGCCATCATCGGCGCGGACCTTCTACGTGGAGCGGCGCGGCAACCG 2015
Db 224 GTTCGAGATCGACGCCAACCGCATCTCTGAGGTGTGCGGCCGAGGACAGGCGCGGCAA 283
QY 2016 CGAGACCAAGATCGCAAGCGCGGTACGTGACCGACCGCGCGCGGCAAGAGATCGTGAG 2075
Db 284 CCGGAGAGATGCTCATCACCACGACCAAGACCGCTGACGCGCGGACGATCGAGCG 343
QY 2076 CCTGACCGAGACCAACACAGAGACCGAGCTGCGAGCCATCCAGCTGGCCCTGCAGGA 2135
Db 344 CATGATCAAGGATCGGAGCGGTTTCGCGGACGACGACCAAGAGCTGAAGGAGCGCGTGA 403

```

```
QY 2136 CAGCGGACGAGGTGAACATCTGTGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGC 2195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 404 GGCCCGCAACGAGCTCAGAGCTACGGCTACAGCTGAAGAACACGAGCTCAGCTCGAAGGA 463
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 2196 CGAGCCCGCAAGAGCGAGCGAGCTGGTGACACCATCATCGAGCAGCTGATCAAGAA 2255
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 CAAGCTGGGCGGAGCGGTGTCCGACGACGACAAAGGCGAAGTGAAGGACGATCGACGA 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2256 GGAGAAGGTGTACTGAGCTGGGTGCCCGCCACAGAGGCATCGGCGGCAACGAGCAGAT 2315
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 GAGATCAAGTGTGCTGACAGAACACAGACACGAGCGCGGAGAGTACAAAGACAGAA 593
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2316 CGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTGTCTCGACGCGCATCGATGCGG 2375
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 GAAGGAGCTGGAGACATCTGCAGCCCATCATTCGCAAGCTGTACGCGAGCAGTGGCGG 643

RESULT 15
CNS0091P      925 bp      DNA      linear      GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION    AL053013
VERSION      AL053013.1 GI:4934461
KEYWORDS     GSS.
SOURCE       Drosophila melanogaster.
ORGANISM     Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE    1 (bases 1 to 925)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
              BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
              collaboration with the Berkeley Drosophila Genome Project (BDGP).
              The BDGP is constructing a physical map of the Drosophila
              melanogaster genome using these BACs. For further information
              please see http://www.fruitfly.org The BDGP Drosophila
              melanogaster BAC library was prepared by Kazutoyo Osoegawa and
              Aaron Mammosier in Pieter de Jong's laboratory in the Department of
              Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
              NY. The library is named RPCI-98 and was constructed by partial
              EcoRI digestion of Drosophila DNA provided by the BDGP from the
              isogenic strain y2; cn bw sp, the same strain used for the BDGP's
              p1 and EST libraries. A more detailed description of the library
              and how to order individual BAC clones, the entire library, or
              filters for hybridization from the BACPAC Resource Center can be
              found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
FEATURES     Location/Qualifiers
              1..925
                /organism="Drosophila melanogaster"
                /db_xref="taxon:7227"
                /clone="BACR19D16"
                /clone_lib="RPCI-98"
                /note="end : TET3"
BASE COUNT   120 a      61 c      61 g      172 t      511 others
ORIGIN
```

Query Match 2.9%; Score 71; DB 17; Length 925;
Best Local Similarity 15.6%; Pred. No. 0.026;
Matches 58; Conservative 173; Mismatches 140; Indels 0; Gaps 0;

```
QY 141 CGCCCGCCGCAAGAGGCTGCTGGAAGTCGCGCAAGGAGGCGCACGATGAAGGACTG 200
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 924 SBSCSCSCSBSBSSSNMTSSNSBSCSSBSSTSSMSSSSSSSGSSSSS 865
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 201 CACCGAGCGCGGACCACTTCTTCGCGAGGACCTGGCCTTCCCGCAGGCAAGCCCG 260
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: February 10, 2003, 20:43:27
Job time : 3412.25 secs

```
Db 864 GTSSACVKCNASSSSCCCGCMABCCMCMSSSSSSCCSSASARGVKYVRASGAGKRGGSG 805
QY 261 CGAGTTCCCGCAGCAGAGCAACCGCGCCACAGCCCGCAGCTGCAGGTGCG 320
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 804 GASNHSSSSACBSSSSCSASCSWSASSSSSRSSRSGGAGCGSASSRSSSSSSA 745
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 321 CGGCGCAACCCCGCAGCAGGCGCGCGCGAGCGCCGAGCGCCAGGCGCACCTGAACCTCCCCCA 380
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 744 SAGSVSSASSSSSSSSSVSCSVASSMSCSBSSSSSASASSSSSSSSSSSASCSCCCT 685
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 381 GATCACCCCTGTGCAGCGCCCTCTGTGTGACATCAAGTGGCGCGCCAGATCAAGGAGGC 440
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 684 SWSCSCTSSASMAARSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSGSACGBS 625
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 441 CCTGCTGGACACCGCGCGCGACACACCGTCTGTGAGGAGATGAGCCTGCCCGCAAGTG 500
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 MSSGGGSGSVASSGMSVSSSGRSGSGGGGSGSSSSSSSSSSSSSGSGSGSVCSGSS 565
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 501 GAAGCCCCAAGA 511
      | : | : |
Db 564 GCMCRCSGSA 554
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:14:58 ; Search time 80.1949 Seconds
(without alignments)

Title: US-09-610-313-30
 Perfect score: 2469
 Sequence: 1 gtcagccaccatgqcca.....gaactagaccggtaaatc 2469

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
Issued_Patents_NA.*
1: /cgn2-6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2-6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2-6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2-6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2-6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgn2-6/ptodata/2/ina/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1132	45.8	2601	4	US-09-117-217-7		Sequence 7, Appli
2	1132	45.8	2601	4	US-09-117-217-9		Sequence 9, Appli
3	1132	45.8	2601	4	US-09-117-217-11		Sequence 11, Appli
4	1132	45.6	2601	4	US-09-117-217-13		Sequence 13, Appli
5	1125.6	45.6	7399	2	US-08-418-848A-9		Sequence 9, Appli
6	1125.6	45.6	9709	2	US-08-188-583-5		Sequence 5, Appli
7	1125.6	45.6	9709	3	US-08-388-353-1		Sequence 1, Appli
8	1125.6	45.6	9709	3	US-08-488-551B-1		Sequence 1, Appli
9	1125.6	45.6	9709	4	US-09-309-572-15		Sequence 15, Appli
10	1125.6	45.6	12494	4	US-08-935-312-13		Sequence 13, Appli
11	1125.6	45.6	12494	4	US-08-848-760B-33		Sequence 33, Appli
12	1125.6	45.6	15581	3	US-08-646-538-35		Sequence 35, Appli
13	1125.6	45.6	15581	4	US-09-503-222-35		Sequence 35, Appli
14	1116	45.2	9737	2	US-08-944-449-7		Sequence 7, Appli
15	1116	45.2	9737	4	US-09-353-362-7		Sequence 7, Appli
16	1080.8	43.8	9746	1	US-08-022-835-3		Sequence 3, Appli
17	1080.8	43.8	9746	1	US-08-388-809-3		Sequence 3, Appli
18	1080.8	43.8	9746	2	US-08-647-714-3		Sequence 3, Appli
19	1079.6	43.7	8932	4	US-09-124-900-1		Sequence 1, Appli
20	1079.6	43.7	8933	3	US-08-463-210-4		Sequence 4, Appli
21	1079.6	43.7	8933	4	US-09-620-958A-3		Sequence 3, Appli
22	1079.6	43.7	8933	4	US-08-620-958A-4		Sequence 4, Appli
23	1079.6	43.7	8933	4	US-09-620-958A-9		Sequence 9, Appli
24	1067.2	43.2	9739	1	US-08-022-835-1		Sequence 1, Appli
25	1067.2	43.2	9739	1	US-08-388-809-1		Sequence 1, Appli
26	1067.2	43.2	9739	2	US-08-647-714-1		Sequence 1, Appli
27	1062	43.0	5362	3	US-08-463-210-5		Sequence 5, Appli


```

1876 CCACAGAAAGCATAGTAATATGGGAAGAACTCCTAAATTTAAACTGCCCATACAAAGG 1935
1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGCGCCACCTGGATCCCGAGTGGAGT 1918
1936 AAACATGGGAACATGGTGACAGAGATATGGCAAGCCACCTGGATTCCTGTGGGAGT 1995
1919 TCGTGAACACCCGCCCTGGTGAAGCTGTGTACCAGCTGGAGAAGGAGCCCATCATCG 1978
1996 TTGTTAATACCCCTCCTTAGTGAAATATGGTACCAGTTAGAGAAGAACCCATAGTAG 2055
1979 GCGCCGAGACCTTCTACGTGGAGCGCGCCGCAACCGCGGAGACCAAGATCGGCAAGCGG 2038
2056 GAGCAGAAACCTTCTATGTAGATGGGCGAGCTAACAGGGAGACTAAATTAGGAAAAGCAG 2115
2039 GCTACCTGACCGACCGCGCGCAAGATCGTAGGCTGACCGGAGACCAACACAGCA 2098
2116 GATATGTTACTAATAGAGGAACAAAAAGTTGTCAACCTTAACACACAAACATCAGA 2175
2099 AGACCGAGTGCAGGCGATCCAGCTGGCCCTGCAGGACACGCGGAGGTGAACATCG 2158
2176 AGACTGAGTTACAAGCAATTTAATCTAGCTTTGCAGGATTTCGGGATTAGAAATAAATAG 2235
2159 TGACCGACAGCAGTAGTACGCCCTGGGCATCATCCAGCGCCAGCCGACACAGAGCGAGCG 2218
2236 TAACAGACTCACAAATATGCATTAGGAATCATTTCAAGCACAAACCCAGATCAAAAGTGAATCAG 2295
2219 AGCTGTGAACACAGATCATCGACAGCTGTATCAAGAGAGGAGAGGTGTACCTGAGCTGG 2278
2296 AGTTAGTCAATCAATTAATAGACAGTGTATAAAAAAGGAAAGGTCTATCTGGCATGGG 2355
2279 TGCCCGCCCAACAGGCGATCGCGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2338
2356 TACCAGCACACAAAGGAATGGAGGAATGAACAAGTAGATAAATAGTCAGTGTCTGGAA 2415
2339 TCAGCAAGGTGCTTCTCGTGGAGCGGATCGAT 2370
2416 TCAGAAAGTACTATTTTAGATGGAATAGAT 2447
```

RESULT 3

```

US-09-117-217-11
; Sequence 11, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (453)..(749)
; OTHER INFORMATION: Protease
US-09-117-217-11
```

```

Query Match 45.8%; Score 1132; DB 4; Length 2601;
Best Local Similarity 68.5%; Pred. No. 8,6e-178;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;

QY 14 TGCCCGGAGCCATGAGCCAGGCCACCA---GCGCCAAACATCCTGATGACGCGCAGCAACT 70
DB 77 TGGCTGAAGCAATGAGCCCAAGTAACAAATTCAGCTACCAATAATGATGCAGAGAGGCAATT 136
```

```

QY 71 TCAGGGGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGCAAGAGAGGCCACATCGCCC 130
DB 137 TTAGGAACCAAAAGAAATGTTAAGTGTTCATTTGTGCAAAAGAGGCGCACACAGCCA 196
QY 131 GCAACTGCGCGCGCCCGCGCAAGAGGCTGTGGAAGTGGCGCAAGAGGCGCCACAGCA 190
DB 197 GAAATTCAGGGGCGCCCTAGGAAAAGGGCTGTGGAATAATGTGAAAAGAGGACACACAAA 256
QY 191 TGAAGACATGCACCGAGCGCCAGCCAACTTCTCCGCGAGGAGACCTGGCCCTTCCCCCAGG 250
DB 257 TGAAGATTGTACTGAGAGACAGGCTAA--TTTTTTAGGGAAGATCTGGCCCTTCTTACAAG 315
QY 251 GCAAGCGCGCGAGTTCCCGCAGCAGCAACCGCGCAACACAGCCCGCCACACAGCGCGAGC 310
DB 316 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACAGACACACGCCCCACACAGAGAGAGC 375
QY 311 TGCAGTGGCGCGG-----CGAACACCCCGCAGCGAGCGCGCGGCGGAGCGCCAGGGCA 364
DB 376 TTCAGGTCTGGGTAGAGACACAACCTCCCGCTCAGAAGCAGGAGCGCATAGACAAGGAA 435
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGCTGAGCATCAAG 418
DB 436 CTGTATCTCTTAACCTTCCCTCAGGTCTACTCTTTGGCAACGACCCCTCGTCAACAATAAGA 495
QY 419 TGGCGCGCCAGATCAAGAGGCGCCTGTGACACACCGCGCGCGAGCACACCGTGTGAGG 478
DB 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGTATTAGAAG 555
QY 479 AGATGAGCTGCGCGGCAAGTGGAAAGCCCAAGATGATCGCGCGGATCGCGCGGCTTCATCA 538
DB 556 AAATGAGTTTGCAGGAAGATGGAACCAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 615
QY 539 AGTGCGCGCAGTAGCAGCAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCG 598
DB 616 AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGGACATAAAGCTATAGGTACAG 675
QY 599 TGCTGATCGGGCCCCACCCCGTGAACATCATCGCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 676 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTTGCACCTCAGATTGGTT 735
QY 659 GCACCTGAACTTCCCGCATCAGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCCCGGCA 718
DB 736 GCACCTTAAATTTTCCCATTTAGCCCTATTGAGACTGTGTACCAAGTAAATTTAAGCCAGGAA 795
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 796 TGGATGCGCCCAAGGTTAAACAATGGCCATTGACAGAAGAAAATAAAGACATTAGTAG 855
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGCGGTGACCCGTGCTGGACGTGGCGGCGAGACCCCT 838
DB 856 AAATTTGTACAGAGATGGAAGAGGAGCGAAATTTCAAAAATTTGGCCCTGAAATTCAT 915
QY 839 ACAACACCCCGTGTTCGCCCATCAAGAAGAGACAGACACCAAGTGGCGGCAAGCTGGTGG 898
DB 916 ACAATCTCCAGTATTTGCCATAAAGAAAAGACAGTACTATAATGGAGAAAATTAGTAG 975
QY 899 ACTTCGCGGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCC 958
DB 976 ATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGCCCTGAAGAAGAAAGAGCGGTGACCCGTGCTGGACGTGGCGGCGAGCCCTACT 1018
DB 1036 ATCCCGCAGGCTTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGTGATGCATATT 1095
QY 1019 TCAGCTGCGCCCTGGAGGAGACTTCGGCAAGTACACCGCTTCACCATCCCGAGCATCA 1078
DB 1096 TTTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATATCTGCAATTTACCATACCTAGTATA 1155
QY 1079 ACAACAGAGACCCCGCATCCGCTACAGTACAACGTTGCTGCCCGAGGCGTGGGAAGGCA 1138
DB 1156 ACAATGAGACACCCAGGATTAGATATCAGTACAAATGTGCTTCCACAGGGATGGAAGGAT 1215
```

QY 1139 GCCCAGCATCTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACC 1198
Db 1216 CACCAGCAATATTCCAAAGTAGCATGACAAAATCTTAGAGCCCTTTAGAAAACAAAATC 1275
QY 1199 CCGAGATCGTGATCTACCAAGTACATGAGCAGCTGTAGTGGGAGCAGCACTGGAGATCG 1258
Db 1276 CAGACATAGTTATCTAATACATGAGTATGTTGTAGTACTGACTAGAAATAG 1335
QY 1259 GCCAGCCGCGCCAAAGATCGAGGAGTGCACCAAGCAGCTCTCGCTGGGGCTTCACCA 1318
Db 1336 GGCAGCATAGAACAAAATAGAGGAGCTGACACACATCTCTTGGAGTGGGACTTACCA 1395
QY 1319 CCCCGCAAGAAGCACCAGAGGAGCCCTTCCTGTGGATGGGTACAGCTGCACC 1378
Db 1396 CACCAGACAAAACATCAGAAAGAACCTCCATTCTTGGATGGGTTATGAACCTCCATC 1455
QY 1379 CCGACAAGTGCACCGTGGAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1438
Db 1456 CTGATAATGGACAGTACAGCCTATAGTCTGCCAGAAAAAGACAGCTGGACTGTCAATG 1515
QY 1439 ACATCCAGAGCTGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGCATCAAG 1498
Db 1516 ACATACAGAAATAGTGGGAAATGAAATGGGCAAGTCAAGTTTACCCAGGGATTAAAG 1575
QY 1499 TGCGCCAGCTGTCAAGCTGTGCGCGCGCCAAAGGCCCTGACCGACATCGTGCCTGTA 1558
Db 1576 TAAGCAATATGTAACTCTTAGAGGAACCAAGACACTAACAGAGTAATACCCTAA 1635
QY 1559 CCGAGAGGCGGAGCTGGAGTGGCCGAGAACCGCGAGATCCTGCGGAGCCCTGTGCAG 1618
Db 1636 CAGAAAGAGCAGAGCTAGAACTGCGAGAAAAACAGAGAGATTCTAAAGAACCACTACATG 1695
QY 1619 GCGTGTACTACGACCCAGCAGAGCACTGGTGGCGGAGATCCAGACAGGCGCACGACC 1678
Db 1696 GAGTGTATTATGCCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 1755
QY 1679 AGTGAGCCTTACCAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 1738
Db 1756 AATGGACATATCAAAATTTATCAAGGCACTTTAAATCTGAAAACAGAAAATATGCA 1815
QY 1739 AGATGGCAGCCGCCACACCAAGCAGTGAAGAGCTGACCGAGCGCGTGCAGAGATCG 1798
Db 1816 GAATGAGGGGTGCCACACTAATGATGTAAACAAATTAACAGAGCGAGTGCAAAATAA 1875
QY 1799 CCATGGAGAGCATCGTGATCTGGGCAAGACCCCAAGTCTCCGCTGCCATCCAGAGG 1858
Db 1876 CCACAGAAAGCATAGTATATATGGGAAAGACTCTTAATTTAACTGGCCATACAAAGG 1935
QY 1859 AGACCTGGGAGACCTGTGTGACCGACTACTGGCAGGCCACCTGGATCCCGGAGTGGAGT 1918
Db 1936 AAACATGSGAAACATGTGGACAGATATTGGCAAGCCACCTGGATTCTCTGAGTGGAGT 1995
QY 1919 TCGTGAACACCCCGCCCTGTGTGAAGCTGTGGTACCGAGTGGGAAGAGGCCCATCATCG 1978
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATATTGTTACCAAGTTAGAGAAAGAACCCCATAGT 2055
QY 1979 GCGCGGAGACCTTCTAGCTGACGCGCGCCGCAACCCGAGACCAAGATCGCAGGCGC 2038
Db 2056 GAGCAGAAACCTTCTATGTAGTGGGAGCTTAACAGGGAGACTAAATTAGGAAGAGAG 2115
QY 2039 GCTACGTGACCGACCGGGCGCGCAGAGATCGTGAAGCTGACCGGAGACCAACCAAGAG 2098
Db 2116 GATATGTTACTAATAGAGGAGACAAAAGTTGTACCCCTAACTGACACACAAATCAGA 2175
QY 2099 AGACCGAGCTCGAGGCATCCAGCTGGCCCTGCGAGGACAGCGGAGCGAGGTGAACATCG 2158
Db 2176 AGACTGAGTTTACAAGCAATTTATCTAGCTTTGCAGGATTTCGGGATTAGAAGTAAACATAG 2235
QY 2159 TGACCGAGACCGACTACGCTGGCATCATCCAGGCCCGCCAGCAGAGGCGAGAGCG 2218
Db 2236 TAAACAGACTCAAAATATGCAATAGCAATGCAATAGCAATCAAGCACAACCAAGATCAAAATCAG 2295
QY 2219 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGGAGTGTACCTGAGCTGGG 2278

Db 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAAGGAAAGGTCTATCTGGCATGGG 2355
QY 2279 TGCCCGCCCAAGGCGATCGCGGCAAGCAGCAGATCGAAGCTGTGAGCAAGGGCA 2338
Db 2356 TACCAGCACACAAAGGAATTGGAGGAATGAACAAAGTAGATAAATTAGTCAGTGTGAA 2415
QY 2339 TCCGCAAGGTGCTGTCTCTGGACGCATCGAT 2370
Db 2416 TCAGAAAGTACTATTTTATGATGAATAGAT 2447

RESULT 4
US-09-117-217-13
; Sequence 13, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOFS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 13
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (750)..(2435)
; OTHER INFORMATION: Reverse Transcriptase
US-09-117-217-13

Query Match 45.8%; Score 1132; DB 4; Length 2601;
Best Local Similarity 68.5%; Pred. No. 8.6e-178;
Matches 1626; Conservative 0; Mismatches 730; Indels 16; Gaps 4;

QY 14 TGGCGGAGCCCATGAGCCAGGCGCACCA---GCGCCAACTCTGATGCGCGCAGCAACT 70
Db 77 TGGCTGAGCAATGAGCCAAAGTACAAATTCAGCTACCAATATGATGAGAGGCAATT 136
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCACTGCGGCAAGGAGGCCACATCGCC 130
Db 137 TTAGGAACCAAGAAAGATTGTTAAGTCTTTCAATTGTGGCAAGAGGGCACACAGCA 196
QY 131 GCAACTGCGCGCCCGCCCGCAGAGAGGGCTGCTGGAAGTGGCGGCAAGAGGCCACACGA 190
Db 197 GAAATTGCGAGGCGCCCTAGGAAAAAGGGCTGTTGGAATGTGGAAGGAAGGACACCAA 256
QY 191 TGAAGGACTGACCGAGCGCCAGGCAACTTCTTCCGCGAGGACCTGCGCTTCCCCCAGG 250
Db 257 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTATGGAAGATCTGGCTTCTCTACAG 315
QY 251 GCAAGGCGCGGAGTTCCTCCAGCGAGAGAACCGCGCAACAGCCCGCCAGCGCGAGC 310
Db 316 GGAAGGCGAGGAAATTTTCTTCAGAGCAGACAGAGCCCAACAGCCCCCACCAGAGAGC 375
QY 311 TGCAGGTGCGGG-----CGACAAACCCCGCAGGAGCGCGCGCCGCGGCGCCAGGCA 364
Db 376 TTCAGGTCTGGGGTAGAGACAACTCCCTCTCAAGAGAGGAGCGCCATAGACAGGAA 435
QY 365 -----CCTGAACCTTCCCCAGATCACCTGTGCGAGCGCCCTGGTGAGCATCAAG 418
Db 436 CTGTATCTTTAACTTCTCCTCAGGTCACTCTTTGGCAACAGCCCTCGTCACAATAAGA 495
QY 419 TGGCGGCGCAGATCAAGGAGGCCCTGCTGGACACCGGGCGCGACACCGCTGTGTGAGG 478
Db 496 TAGGGGCGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATCATACAGTATTAGAG 555

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/418,848A
FILING DATE: 07-APR-1995
CLASSIFICATION: 526
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936,849
FILING DATE: 28-AUG-1992
CLASSIFICATION: 526
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-4091US2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7399 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-418-848A-9

Query Match 45.68; Score 1125.6; DB 2; Length 7399;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGGCGAGGCCATGAGCGAGCGCCACCA---CGGCCAACATCTGTATGAGCGCGAGCAACT 70
DB 1838 TGGCTGAAGCAATGAGCGCAAGTACAAATCCAGCTACCATTATATATACAGAAAGGCAATT 1897

QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGGCGAAGAGGGGCCACATCGGCC 130
DB 1898 TTAGGAACCAAGAAGAGACTTTAAGTGTCTTCAATTTGGCANAAGAGGGGCACATAGCCA 1957

QY 131 GCAACTGCGCGCCCCCGGCAAGAGGGCTCTGGAAGTGGCGGCAAGAGGGCCACACAGA 190
DB 1958 AAAATTGAGGGCCCCCTAGGAAAAGGGCTGTGGAAATGTGGAAGAGGAGACACAAA 2017

QY 191 TGAAGAGTGCACCGAGCGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGG 250
DB 2018 TGAAGAGTTGTACTGAGAGACAGCTAA-TTTTGTAGGAGATCTGGCCCTTCCCAAG 2076

QY 251 GCAAGGCCCGGAGTTCCCGACGAGCAGAGAACCGCGCCACACAGCCCAAGCGCGAGC 310
DB 2077 GGAAGGCCAGGGAAATTTTCTTCAGAGCAGACACAGAGCCCAAGCCCAAGAGAGAGC 2136

QY 311 TGCAGGTTGCGGG---CGACAAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCGCA 364
DB 2137 TTCAAGTTTGGGGAAGAGACAACAACTCCCTCTCAGAGAGGAGGCGCGGATAGACAAAGGAA 2196

QY 365 -----CCCTGAACCTTCCCGACATCACTCTGTGGCAGCGCCCTGTGTAGCATCAAGG 418
DB 2197 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCCACCCCTCGTCACATAAGA 2256

QY 419 TGGCGCGCAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGCACCCCTGTGTGAGG 478
DB 2257 TAGGGGGCAATTAAGAGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2316

QY 479 AGATGACCTTCCCGGCAAGTGAAGCCCAAGATGTCGGCGCATCGCGGCTTCATCA 538
DB 2317 AATGAATTTGCCAGGAGATGGAACCAAAATGATAGGGGAAATTTGGAGGTTTTATCA 2376

QY 539 AGGTGCCCGAGTACGACCATCTGTATCGAGATCTCGGCAAGAGGCCATCGGCACCG 598
DB 2377 AAGTAGGACAGTATGATCAGATCTATAGAAATCTCGGACATTAAGCTATAGGTACAG 2436

QY 599 TGCTGATCGGCCCCACCCCGTGAACATCATCGCGGCAACATGCTGACCCAGCTGGGCT 558
DB 2437 TATTAGTAGACCTACACTGTCAACATAATTGGAGAATAATCTGTGACTCAGATTGGCT 2496

QY 659 GCACCTTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCA 718
DB 2497 GCACCTTAAATTTTCCCATATTAGTCTATTGAGACTGTACCACTMAAATTAAGCCAGAA 2556

QY 719 TGGAGGGCCCCAAGGTGAAGCAGTGGCCCTGACGAGGAGAGATCAAGGCGCTGACCG 778
DB 2557 TGGATGGGCCCAAAAGTTAAACAAATGGCCATTGACAGAGAAAGAAAATAAAGCAATTAGTAG 2616

QY 779 CCATCTGCGAGGAGTGGAGAAGAGGAGGCAAGATCACCAAGATCGSCCCCGAGACCCCT 838
DB 2617 AATTTGTACAGAAATGGAAGAGGAGGAAATTTTCAAAATTTGGCCCTGAAATTCAT 2676

QY 839 ACAACACCCCGGTGTTCCGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
DB 2677 ACAATACTCCAGTATTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAATTAGTAG 2736

QY 899 ACTTCCGCGAGCTGAACAAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCCC 958
DB 2737 ATTTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCA 2796

QY 959 ACCCGCGCGCTGAAAGAAGAGAGCGTGAACCGTGTGACGCTGGCGAGCGCTACT 1018
DB 2797 ATCTGTCAGGTTAAACAGAAAAAATCAGTAAACAGTACTGGATGTGGCGATGATATT 2856

QY 1019 TCAGGCTGCCCTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCTCCCGACATCA 1078
DB 2857 TTTCACTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCAATTTACCATACCTAGTATAA 2916

QY 1079 ACAGGACACCCCGGCTACCGTACCAAGTACAAAGTGTGCTGCCCGAGGCTGGAGGCA 1138
DB 2917 ACAATGAGACACCGGATATAGATATAGTACAATGCTTCCACAGGATGGAAAGAT 2976

QY 1139 GCGCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGCGCGCCCAACC 1198
DB 2977 CACCAGCAATATTCAGTGTAGCATGACAAATCTTAGAGCCTTTTGAANAACAAATC 3036

QY 1199 CCGAGATCGTATCTACAGTACATGAGCAGCTGTAGTGGCGAGCAGCTGGAGATCG 1258
DB 3037 CAGACATAGTCTATCAATACATGAGTATTTGTATGTAGGATCTGACTTAGAATAAG 3096

QY 1259 GCGAGCAGCGCGCAAGATCGAGGAGTGCAGCAGCACTGCTGCTGCTGGGCTTACCA 1318
DB 3097 GGCAGATAGAACAAAATAGAGGAACCTGAGACACATCTGTTGAGGTGGGATTTACCA 3156

QY 1319 CCGCCGACAAAGAACACAGAGGAGCGCCCTTCCCTGTGGATGGCTACGAGCTGCACC 1378
DB 3157 CACCAGACAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGTTTATGAACTCCATC 3216

QY 1379 CCGACAGTGAACCGTGCAGCCCATCGAGCTGCGCGAGAGGAGAGCTGGACCGTGAACG 1438
DB 3217 CTGATAATGACAGTACAGCTTATAGTGTGCGCAAGAACAGCAGCTGACTGTCAATG 3276

QY 1439 ACATCCAGAACTGCTGGGCAAGCTGAACCTGGCGCAGCAGCATCTACCCCGCATCAGG 1498
DB 3277 ACATACAGAAATTAGTGGGAAAATTTGAATTTGGGCAAGTCAAGTTTATCGAGGATTAAG 3336

QY 1499 TGCCTCAGCTGTGAAGCTGCTGCGCGCGCGCAAGCGCCCTGACCCACATCTGTCGCCCTGA 1558
DB 3337 TAAGGCAATTATGATAAATCTTAGGGGAACCAAGCACTAACAGAGTATGTAACCTAA 3396

QY 1559 CCGAGGAGCGGAGCTGGAGTGGCGGAGAACCGCGAGATCCTGCGGAGCGCCCTGACG 1618
DB 3397 CAGAAGAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTTCTAAAAGAACCCGTTACATG 3456

QY 1619 GCGTGTACTAGACCCCGCAGCAAGCACTGCTGGCGGAGATCCAGAGCAGGCGCCACGACC 1678
DB 3457 GAGTGTATTATGACCCCATCAAGAGACTTAATAGCAGAAATACAGAAAGCGGGCAAGGCC 3516

QY 1679 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCA 1738

Db	3517	ATATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTCAAAACAGAAAAATATGCAA	3576
QY	1739	AGATGGCAGCCGCACACCAACAGCGTGAAGCAGCTACCGAGGCGCTCCAGAAAGATCG	1798
Db	3577	GAATGAAGGTCGCCACACTATATGATGTAACAATTAACAGAGGCAGTACAAAAATAG	3636
QY	1799	CCATGGAGAGCATCGTATCTGGGGCAAGACCCCCCAAGTTCCGCTGCCCATCCAGAAG	1858
Db	3637	CCACAGAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAAATTACCCATACAAAGG	3696
QY	1859	AGACCTGGGAGACCTGTGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT	1918
Db	3697	AAACATGGGAAGCATCGTGGACAGATATTGGCAAGCCACCTGGATTTCCTGACTGGGAGT	3756
QY	1919	TCGTGAACACCCCCCTGTGTGAAGCTGTGTACCAAGCTGGAGGAGGCCCATCATCG	1978
Db	3757	TTGTCTAATACCCCTCCCTTAGTGAAGTTATGGTACCAAGTTAGAGAAAGACCCATATAG	3816
QY	1979	GGCCCGAGACCTTCTACGTGGACGGCGCGCCAAACCGCGAGACCAAGATCGGCAAGCCG	2038
Db	3817	GAGCAAAACTTCTATGTATAGATGGGCAGCCAAATAGGAAACTAAATTAGGAAAAGCAG	3876
QY	2039	GCTACGTGACGACCGGGCGCGCAAGAAGATCTGTAGCGTGACCGAGCACCAACACCAGA	2098
Db	3877	GATATGTAATGACAGAGGAAGACAAAAAGTTGTCCCTTAACGGACACAAACAAATCAGA	3936
QY	2099	AGACCGAGCTGAGAGCCATCCAGCTGGCCCTGCAGGACAGCGGCAGCGAGTGAAACATCG	2158
Db	3937	AGACTGAGTTTACAGCAATTCATCTAGCTTTGCAGAGTTTCGGGAAATAGAAGTAACATAG	3996
QY	2159	TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACGAAGAGCGAGAGCG	2218
Db	3997	TGACAGACTCACAATATGCATTGGGAATCATTCAGACACACCCAGATAAGAGTGAATCAG	4056
QY	2219	AGCTGTTGAACAGATCATCGAGCAGCTGATCAGAGGAGGAAGTGTACCTGAGCTGGG	2278
Db	4057	AGTTAGTCAGTCAAAATAATAGAGCAGTTTAATAAANAAGGAAAAAGTCTACTCTGGCATGGG	4116
QY	2279	TGCCCCCCACAAAGGCGATCGCGGCCAACGAGCAGATCGCAACGCTGGTGAGCAAGGCA	2338
Db	4117	TACCACACACAAAGGAATTGGAGGAATGACACAGTAGATGGGTTGGTCAGTCTGGAA	4176
QY	2339	TCGCAAGGTGCTGTTTCCTGGAGCGCATCGAT	2370
Db	4177	TCAGGAAATCTACTATTTTATAGATGAATAGAT	4208

RESULT 6

```

1  RESULT 0
2
3  US-08-188-583-5
4  : Sequence 5, Application US/08188583
5  : Patent No. 5851813
6
7  GENERAL INFORMATION:
8
9  : APPLICANT: Destosiers, Ronald C.
10 :
11 : TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES
12 :
13 : NUMBER OF SEQUENCES: 57
14 :
15 : CORRESPONDENCE ADDRESS:
16 :
17 : ADDRESSEE: Fish & Richardson
18 :
19 : STREET: 225 Franklin Street
20 :
21 : CITY: Boston
22 :
23 : STATE: Massachusetts
24 :
25 : COUNTRY: U.S.A.
26 :
27 : ZIP: 02110-2804
28
29 : COMPUTER READABLE FORM:
30 :
31 : MEDIUM TYPE: 3.5" Diskette, 1.44 MB
32 :
33 : COMPUTER: IBM PS/2 Model 50Z or 55SX
34 :
35 : OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
36 :
37 : SOFTWARE: WordPerfect (Version 5.0)
38
39 : CURRENT APPLICATION DATA:
40 :
41 : APPLICATION NUMBER: US/08/188,583
42 :
43 : FILING DATE:
44 :
45 : CLASSIFICATION: 435
46 :
47 : PRIOR APPLICATION DATA:
48 :
49 :
50 :
51 :
52 :
53 :
54 :
55 :
56 :
57 :
58 :
59 :
60 :
61 :
62 :
63 :
64 :
65 :
66 :
67 :
68 :
69 :
70 :
71 :
72 :
73 :
74 :
75 :
76 :
77 :
78 :
79 :
80 :
81 :
82 :
83 :
84 :
85 :
86 :
87 :
88 :
89 :
90 :
91 :
92 :
93 :
94 :
95 :
96 :
97 :
98 :
99 :

```

QY 719 TGGACGGCCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGGATCAAGGCCCTGACCG 778
DB 2596 TGGATGGGCCAAAAGTTAAACAATGGCCATTGACAGAGAGAAAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTGCGAGGAGATGAGAGAGGAGGCAAGATCACCAGATCGGCCCGCCGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGGAAGAGAGGAGAAATTTCAAAAATTTGGGCTGAAAATCCAT 2715
QY 839 ACAACACCCCGCTGTCGCCATCAAGAGAGGACAGCACCAAGTGGCGCAAGCTGGTG 898
DB 2716 ACAATCTCCAGTATTGTCCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2775
QY 899 ACTTCGCGGAGCTGAACAAGCCACCCAGGACTTCTGGAGGTGCGAGCTGGCGATCCCCC 958
DB 2776 ATTTACAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAATTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGCTGAAG 1018
DB 2836 ATCTGCGAGGTTAAACAGAGAAATTCAGTAACAGTACTGGATGTGGGCGATGCATATT 2895
QY 1019 TCAGCGTCCCGCTGGAGAGAGACTTCCGCAAGTACACCGCCCTTACCAATCCCGAGCATCA 1078
DB 2896 TTTCAGTCCCTTAGATAAAGACTTCAGGAAGTATCTGCAATTTACCATCTAGTATAA 2955
QY 1079 ACACGAGACCCCGGATCCCGTACAGTACAGTACAGTGTGCTGCCCGAGGGCTGGAAGGCA 1138
DB 2956 ACAATGAGACACAGGATTTAGATATCAGTACAAATCTTGTGAGCCCTTTAGAAAAAATC 3015
QY 1139 GCCCCAGCATCTCCAGAGCAGTACCAAGATCTTGGAGCCCTTCCGCGCCCGCAACC 1198
DB 3016 CACCAGCAATATCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAAATC 3075
QY 1199 CCGAGATCGTGATACCAAGTACAGGACCTGTGAGTGGCGAGGACCTGGAGATCG 1258
DB 3076 CAGACATAGTCATCTCAATACATGATGATTTGTAAGTATCTGATTAAGATAG 3135
QY 1259 GCGAGCAGCGCCCAAGATCGAGAGCTGCGCAGAGCTGCTGCGCTGGGGCTTACCA 1318
DB 3136 GCGAGCATAGAACAAAAATAGAGAACTGAGACAACTCTGTTGAGGTGGGGATTTACCA 3195
QY 1319 CCCCCAGCAAGACAGAGAGGAGGCCCCCTTCTGTTGATGGCTACAGCTGCACC 1378
DB 3196 CACCAGACAAAAACATCAGAGAACTCCATCTTGTGATGGTTATGAATCCATC 3255
QY 1379 CCGCAAGTGGACGTCGACCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACG 1438
DB 3256 CTGATAAATGACAGTACAGCTATAGTGTGCGCAGAAAGACAGCTGGACTGTCAATG 3315
QY 1439 ACATCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGAGGAGATCTACCCGGCATCAAG 1498
DB 3316 ACATACAGAAATAGTGGGAAAAATTTGAATTTGGGCAAGTCAAGATTTATGAGGGATTAAG 3375
QY 1499 TGGCCAGCTGTCAAGCTGCTGCGCGCGCCAGGCCCCCTGACGACATCGTCCCTGA 1558
DB 3376 TAGGCAATATGTAACTTTAGGGGAAACCAAGCACTAACAGAGTAGTACCACTAA 3435
QY 1559 CCGAGAGCGCGCTGGAGCTGGCCGAGAACCCGAGATCTGCGGAGCGCCGTGCAAG 1618
DB 3436 CAGAAGAGCAGAGCTAGAACTGCGAGAAACAGGGAGATTTAAAGAACCCGTACATG 3495
QY 1619 GCGTGTACTAGGACCCAGCAAGGACCTGTGGCGGAGATCCAGAACCGGCCAGCAC 1678
DB 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAAAATACAGAAAGCGGGCAAGGCC 3555
QY 1679 AGTGGACCTACAGATCTACGAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACCCA 1738
DB 3556 AATGGACATCAAAATTTATCAGAGCCATTTAAATCTGAAACAGGAAAAATATGCAA 3615
QY 1739 AGATGGCACCCCAACCAAGCAGTGAAGCAGCTGACCGGCGCTGCAAGAGATCG 1798
DB 3616 GAATGAAGGTGCCCCACATTAATGATGAAACAATTAACAGAGGAGTCAAAAAATAG 3675
QY 1799 CCATGGAGACATCGTGTCTGGGCAAGACCCCAAGTTCCCGCTGCGCATCCAGAGG 1858

DB 3676 CCACAGAAAGCATAGTATATATGGGAAGAGACTCTTAATTTAAATTAACCATACAAAAG 3735
QY 1859 AGACTGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
DB 3736 AAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTAGTGGGAGT 3795
QY 1919 TCGTGAACACCCCGCTGGTGAAGCTGTGTACCAAGTGGAGAGAGCCCATCATCG 1978
DB 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACAGTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCGAGACCTTCTACGTGGACGCGCCCAACCGGAGACCAAGATCGGCAAGGCCG 2038
DB 3856 GAGCAGAACTTTCTATCTAGATGGGCGAGCCAAATAGGAAACTAAATTAGGAAAGCAG 3915
QY 2039 GCTAGTGACCGACCGGGCGCGGAGAGATGTGTAGGCTGACCGAGACCCACCAACAGA 2098
DB 3916 GATATGTAACGAGAGAGAAAGTGTCCCTTAACGGACACAACAAATCAGA 3975
QY 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTCCAGGACGCGGAGGAGGTGAACATCG 2158
DB 3976 AGACTGATTTACAGCAATTCATCTAGCTTTGCGAGGATTCGGGATTAAGATTAACATAG 4035
QY 2159 TGACCGACAGCAGTACGCTTGGCATCTCCAGCGCCAGCCGACAGAGAGAGAGCG 2218
DB 4036 TGACAGACTCAATATGCAATTTGGGAATCATTCAGACACACCAAGATAGAGTGAATCAG 4095
QY 2219 AGCTGGTGAACAGATCATCGAGCAGCTGTATCAAGAGAGAGAGGTGTACCTGAGCTGG 2278
DB 4096 AGTTAGTCAGTCAATAATAGAGCAGTTAATAAAAGGAAAGAAAGTCTACCTGGCATGG 4155
QY 2279 TGCCCGCCCAAGGGGCTGCGGCGCAACGAGCAGATCGACAGCTGCTGAGCAAGGCA 2338
DB 4156 TACCAGCACACAGGAATTTGAGGAAATGAACAGTATAGTGGTGTGCTGCTGCGAA 4215
QY 2339 TCCGCAAGGTGCTGTCTCTGACGCGCATCGAT 2370
DB 4216 TCAGAAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 7

US-08-388-353-1
; Sequence 1, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: McPhee, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9709 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 45.6%; Score 1125.6; DB 3; Length 9709;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGCCGAGGCGCATGAGCCAGGCGACCA---GCGCCAAACATCTCTGACGCGCAGCAACT 70
DB 1877 TGCGTGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCATAATGATACAGAAAGCAATT 1936

QY 71 TCAAGGCGCCCAAGCGGATCATCAAGTCTTCAACTGCGCGCAGGAGGCCACATCGCCC 130
DB 1937 TTAGGAACCAAGAAAGAGCTGTAAAGTGTTCAAATTGTGCAAGAAAGGCGACATAGCCA 1996

QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGCTGTGGAAGTGCAGGAGGCGCACCAGA 190
DB 1997 AAAATTGACGGGCGCCCTAGGAAAGAGGCTGTGGAATGTGGAAGAGGAGACACCAA 2056

QY 191 TGAAGGACTGACCGGCGCCAGGCGCAACTTCTTCCGCGAGGACCTTGGCCCTTCCCCAGG 250
DB 2057 TGAAGGACTGACCGGCGCCAGGCGCAACTTCTTCCGCGAGGACCTTGGCCCTTCCCCAGG 2115

QY 251 GCAAGGCGCCGAGTTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 310
DB 2116 GGAAGGCGCAGGGAATTTCTTTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2175

QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCA 364
DB 2176 TGCAGGTGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2235

QY 365 -----CCCTGAATTCCTCCCGAGATCACTGCTGGCAGCGCCCGCTTGGTGAGCATCAAGG 418
DB 2236 CTGTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2295

QY 419 TGGCGGCGCAGATCAAGGAGGCGCTGCTGACACCGCGCGCGCGCGCGCGCGCGCGCGG 478
DB 2296 TAGGGGGCAATTAAGGAGAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAG 2355

QY 479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGGAAGATGAAACCAAAATGATAGGGGAAATTGGAGGTTTATCA 2415

QY 539 AGGTGCCAGTAGTGGACCATCTGATTCAGATCTGCGGCAAGAGGCGCATCGGCACCG 598
DB 2416 AAGTAGGACAGTATGATCAGATCTACTAGAAATCTCGGACACATAAAGCTATAGGTACAG 2475

QY 599 TGCTGATCGCGCCCGCCCGTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 658
DB 2476 TATTAGTAGACCTACACCTGTCAACATATTTGGAAGAAATCTGTTGACTCAGATTGGCT 2535

QY 659 GCACCTTGAATTTCCCATCAGCCCCATCGAGACCGGTGCGCGCTGGAAGCTGAAGCGCCGCA 718
DB 2536 GCATTTTAAATTTTCCCATTTAGTCTATTGAGACTGTACCAGTAAATTAAGCCAGGAA 2595

QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 2596 TGGATGCCCAAAAGTTTAAACAATGGCCATTGACAGAGGAGGAGGAGGAGGAGGAGGAGG 2655

QY 779 CCATCTCGGAGAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838
DB 2656 AAATTGTACAGAAATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2715

QY 839 ACAACACCCCGTGTTCGCCATCAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898

DB 2716 ACAATACTCCAGTATTGTCATAAAGAAAGACAGTACTATAATGGAGAAATTAGTAG 2775

QY 899 ACTTCGCGAGCTGAACAAGCGCCAGGACTTCTGGGAGGTGCGAGCTGGGATCCCC 958

DB 2776 ATTTTCAGAAACTTAATAAGAGAACTCAAGATTCTGGGAAGTTCAATTTAGGAATACCAC 2835

QY 959 ACCCGCGCGCCTGAAGAAGAAGAGCGTGACCGTGTGGACGTGGGGGCGAGCGCTACT 1018

DB 2836 ATCTCGCAGGGTTAAACACAGAAATAATCAGTAACAGTACTGGATGTGGCGCATGCATATT 2895

QY 1019 TCAGCCTGCGCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078

DB 2896 TTTTCAGTTCCTTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAA 2955

QY 1079 ACAACGAGACCCCGGATCCGCTACAGTACACGCTGTGCTGCCCCAGGCTGGAAGGGCA 1138

DB 2956 ACAATGAGACACCGAGGATTTAGATATCAGTACAAATGTGCTTCCACAGGATGGAAGGAT 3015

QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTCCGCGCGCCGAACC 1198

DB 3016 CACAGCAATATTCAGTGTAGCATGACAAATCTTAGAGCCTTTTAGAAAACAAATC 3075

QY 1199 CCGAGATCTGTGATCTACCATGATACATGAGACACCTGTACGTGGCAGCGACCTGGAGATCG 1258

DB 3076 CAGCATAGTCTATCTATCAATACATGATGATTGTATGTAGGATCTGACTTAGAAATAG 3135

QY 1259 GCAGACACCGCGCCAGAGTGCAGAGCTGCGCAAGCACTGCTGCGCTGGGGCTTCACCA 1318

DB 3136 GCGAGCATAGAACAAATAATAGAGGAACTGAGACAACATCTGTTGAGGTGGGATTTACCA 3195

QY 1319 CCGCCGACAAAGAGCACCAGAGGAGCGCCCTCTCTGATGGATGGGTAGGAGTGCACC 1378

DB 3196 CACGACAAACAAACATCAGAAAGACCTCCATCTCTTGGATGGGTATGAATCTCATC 3255

QY 1379 CCGACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGGAGAGGAGAGTGGACCGTGAACG 1438

DB 3256 CTGATAAATGGACAGTACAGCCTATAGTGTGCGCAAGAAAGACAGCTGGACTGTCAATG 3315

QY 1439 ACATCCAGAGCTGGTGGCAAGCTGAACCTGGCCAGCAGCTACCTCCCGCATCAAGG 1498

DB 3316 ACATACAGAAATTAGTGGGAAATTTGAATTTGGGCAAGTCAAGTTTTCGAGGGATTAAG 3375

QY 1499 TGGCGCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTCACCGACATCGTCCCCCTGA 1558

DB 3376 TAAGGCAATTTATGTAAACTTCTTAGGGGAAACCAAGCACTAACAGAAAGTAGTACCCTAA 3435

QY 1559 CCGAGAGGCGGAGCTGGAGCTGGCCCGAGAACCGCGAGATCTCTGCGGAGGCCCTGCACG 1618

DB 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAACAGGGAGATTCTAAAGAGACCGGTACATG 3495

QY 1619 GCGTGTACTACGACCCAGCAAGGACCTGTTGGCGGAGATCCAGAAAGCAGGGCCAGCACC 1678

DB 3496 GAGTGTATTATGACCCATCAAGACTTATAGCAGAAATACAGAAAGCAGGGGCAAGGCC 3555

QY 1679 AGTGGACCTACCATGATCTACCGAGGCGCTTCAAGAACCTTGAAGACCGCGCAAGTACGCCA 1738

DB 3556 AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGGAAAAATATGCA 3615

QY 1739 AGATGCCACCGGCCACCAACGACGCTGAAGCAGCTGACCGGAGGCCGTGCAGAAAGATCG 1798

DB 3616 GAATGAAGGGTGGCCACACTAATGATGTAAACAAATTAACAGAGGCGAGTACAAAAATAG 3675

QY 1799 CCATGGAGACATCTGATCTGCGGCAAGACCCCAAGTTCCCGCTGCCCTCCCAATCCAGAGG 1858

DB 3676 CCACAGAAACGATAGTAAATTTGGGAAAGACTCCTAAATTTAAATTTACCCATACAAAGG 3735

QY 1859 AGACTGGGAGACTGTTGGACCGACTACTTGGCAGGCCACTCGATCCCGAGTGGGAGT 1918

DB 3736 AAACATGGGAAGCATGTTGGACAGAGTATTGGCAAGCCACCTGGATTCCTTGGTGGGAGT 3795

QY 1919 TCGTGAACACCCCGCTGTTGAGCTGTTGTTACCGACTGGAAGAGGAGGAGGAGGAGGAGGAGG 1978

Db 3796 TTGTCATAACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCCGAGACCTTCTACGTGGAGCGCGCCGCAACCGCAGACCAAGATCGGCAAGCGG 2038
Db 3856 GAGCAAAACTTTCTATGTAGATGGGGCAGCCCAATAGGGAATAAATAGGAAAGCAG 3915
QY 2039 GTTACGTGACCGACCGCGCGGCGAGAAAGATCGTGAGCTGACCGAGACCAACACAGA 2098
Db 3916 GATATGTAACGTACAGAGGAGCAAAAGTTGTCCCTTAAGGGACACAAACAATCAGA 3975
QY 2099 AGACCGAGTGCAGGCGCATTCAGCTGGCCCTGCAGGACAGCGCAGGAGTGAACATCG 2158
Db 3976 AGACTGAGTCAAGCAATTCATCTAGCTTTGCAGGATTCGGGATTAGAAGTAACATAG 4035
QY 2159 TGACCCAGCAGGAGTACGCCCTGGGCGATCATCCAGCGCCGCGCAAGAGGAGCGG 2218
Db 4036 TGACAGCTACCAATATGATTTGGGAATCATTCACGACCAACCAAGATAGAGTGAATCAG 4095
QY 2219 AGCTGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAAAGTGTACCTGAGCTGGG 2278
Db 4096 AGTTAGTCAGTCAATAATAGAGCAGTTAATAAAGGAAAGTCTACCTGGCATGGG 4155
QY 2279 TGCCCGCCCAAGGCGATCGCGCGCAAGCAGCAGATCGACAGCTGGTGAGCAAGGGCA 2338
Db 4156 TACCAGCACACAAGGAATTTGGAGGAATGAACAAGTAGATGGGTTGGTCAGTGCTGGAA 4215
QY 2339 TCCGCAAGGTGCTGTTCCCTGGAGCGCATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTATAGTGGATAGAT 4247

RESULT 8

US-08-488-551B-1
; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: PM4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: PM0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: PM3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGIGLIO
; REFERENCE/DOCKET NUMBER: 9606Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-488-551B-1

Query Match 45.6%; Score 1125.6; DB 3; Length 9709;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGCCCGAGGCGCATGAGCGAGCCACCA---GCGCCAAACATCTGATCGAGCGGCAAGCT 70
Db 1877 TGCTGAGCAATGAGCCAAAGTACAAATCCAGTACCATAATGATACAGAAAGCAATTT 1936

QY 71 TCAAGGCGCCCAAGCGGATCATCAAGTGTTCAACTCGGCAAGGAGGGCCACATCGCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTCAAAGAAAGGCGACATAGCCA 1996

QY 131 GCAACTGCCCGCGCCCGCAAGAGGCTGCTGGAAGTCGGCAAGGAGGGCCACACAGA 190
Db 1997 AAATTTGCGAGGGCCCTAGGAAAGAGGCTGTTGGAATGTGGAAGAGGAGGACACCAA 2056

QY 191 TGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGG 250
Db 2057 TGAAGATTGCTACTGAGAGACAGGCTAA-TTTTATGGGAAGATCTGCGCTTCCCACAAG 2115

QY 251 GCAAGCGCCCGGAGTTCCCGAGCGAGCAGAACCGCGCCACAGCCCGCCACAGCGCGAGC 310
Db 2116 GGAAGCGCAGGGAATTTCTTTCAGAGCAGACAGAGCCACAGCGCCCGCCACAGAGAGGC 2175

QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGAGCGCGAGGCA 364
Db 2176 TTAGGTTGGGGAAGAGACAACACTCCCTCTCAGAAAGGAGGAGCGATAGACAAGGAA 2235

QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGCGAGCGCCCTGGTGAGCATCAAG 418
Db 2236 CTGTATCTTTAGCTTCCCTCAGATCACTTTGGCAGCGACCCCTGTCACATAAAGA 2295

QY 419 TGGGCGCCAGATCAAGAGGCGCCTGCTGGACACCGCGCGCGAGACCGGTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGAGGAGAGATGATACAGATTAGAAG 2355

QY 479 AGATGACCTGCCCGCAAGTGGNAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGGAATTTGGAGGTTTATCA 2415

QY 539 AGGTGCCCGCAGTACGACAGATCCTGATCGAGATCTTGGGCAAGAGGCCATCGGCACCG 598
Db 2416 AGTAGGACAGATGATGATCAGATCACTATAGAAATCTCGGACATATAAGCTATAGTACAG 2475

QY 599 TGCTGATCGGCGCCCGACCCCGTGAACATCATCGCGCCGCAACATGCTGACCCAGTGGGCT 658
Db 2476 TATTAGTAGGAGCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535

QY 659 GCACCTTGAACCTTCCCGCATGAGCCCGATCGAGACCGTCCCGTGAAGCTGAGCGCCGCA 718
Db 2536 GCACCTTAAATTTTCCCATTTAGTCTTATGAGACTGTACCAAGTAAATTAAGCCAGGAA 2595

QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTGTACCGGAGGAGGAAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGCGCCCAAGTTAAACAATGGCCATTGACAGAGAAGAAATAAACCATTAGTAG 2655

QY 779 CCATCTCGAGGAGATGGAGAGGAGGCAAGATCAACAAGATCGGCCCGCGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGAGGAGGAAATTTCAAAAATTTGGGCTGAAAATCCAT 2715

QY 839 ACACACCCCGGTTTCCGCATCAAGAGAGGAGCAGCACCAAGTGGCGCAAGCTGGTG 898
Db 2716 ACAATCTCCAGTATTTGCCATATAAGAAAGACAGTACTAAATGGGAGAAATTAGTAG 2775

QY 899 ACTTCCGCGAGCTGAACAAGCCAGCAGGACTTCTGGGAGGTGACGCTGGGCATCCGCC 958
Db 2776 ATTTCAGAGAACTTAATAAGAGAACTCAAGATTCTCGGAAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGGCTGAAGAAGAAAGAGAGCTGACCGGTCTGGAGGTGGGGGACGCGCTACT 1018
Db 2836 ATCCTGCGAGGGTTAAACACAGAAATAATCAGTAACAGTACTGGATGTGGGGGATGCATATT 2895
QY 1019 TCAGCGTGCCTCGGAGGAGCTTCCGCAAGTACACCGCTTCAACCATCCCGAGCATCA 1078
Db 2896 TTTTCAGTTCCTTAGATAAAGACTTCAGGAAGTATATCTGCTATACCATCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGCGATCCGCTACAGTACAAAGCTGTGCCCCAGGGGTGGAAGGCA 1138
Db 2956 ACAATGAGACACCGAGGATTAGATATCAGTACAAATGTGCTTCCACAGGGATGGAAGGAT 3015
QY 1139 GCCCGAGCATCTTCCAGAGCAGCATGACCAAGATCTCGAGGCCCTTCCGCGCCCGCAACC 1198
Db 3016 CACCAGCAATATTCCAGTGTAGCATGACAAAATCTTTAGAGCCTTTTAGAAAACAAATC 3075
QY 1199 CCGAGATCGTGATCTACCACTACATGGACGACCTGTACGTGGCAGCGACCTGGAGATCG 1258
Db 3076 CAGCATAGTCACTCATCATACATGATGATTTGTATGTAGATCTGACTTGAATAATAG 3135
QY 1259 GCCAGACCCCGGCGAAGATCGAGAGCTGCGCAAGCAACCTGTGCGCTGGGGTTCACCA 1318
Db 3136 GCGAGCATAGAAAATAAGAGAACTGAGACAACATCTGTTGAGGTGGGATTTACCA 3195
QY 1319 CCCCAGCAAGAGCACCAGAGGAGGCCCTTCTCTGTGGATGGGCTAGCAGCTGGACC 1378
Db 3196 CACGACAAAACATCAAGAAAGACCTCCATTCCTTTGGATGGGTATGAACTCCATC 3255
QY 1379 CCGACAAGTGGACGCTGACGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1438
Db 3256 CTGATAATGAGCAGTACACCTTATAGTCTGCCAGAAAGACAGCTGACCTGTCAATG 3315
QY 1439 ACATCAGAAAGCTGTGGGCAAGCTGAACCTGGGCGCAGCCAGATCTTACCCCGGATCAAG 1498
Db 3316 ACATCAGAAATTAAGTGGGAAAATTTGAATTTGGGCAAGTCAGATTTATGCGAGGATTAAG 3375
QY 1499 TGGCCAGCTGTGCAAGCTCTCGCGGCGCCCAAGCCCTGACCGACATCGTCCCGCTGA 1558
Db 3376 TAAGGCAATTAATGTAACCTCTTAGGGGAACCAAGACATTAACAGAAAGTAGTACCACATA 3435
QY 1559 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCCGAGACCTCTGCGGAGCCCGCTGCACG 1618
Db 3436 CAGAAGAAGCAGAGCTAGAAGCTGGCAGAAAACAGGGAGATTTCTAAAGAAACCGGTACATG 3495
QY 1619 GCGTGTACTACGACCCAGCAGGAGCTGGTGGCCGAGATCCAGAGCAGGCGCAGCAC 1678
Db 3496 GAGTGTATTATGACCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGCGCAAGGCC 3555
QY 1679 AGTGGACCTACAGATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCA 1738
Db 3556 AATGGACATATCAAAATTTATCAGAGCCATTTAAAATCTGAAAACAGGAAAATATGCAA 3615
QY 1739 AGATGCGCACCCCGCCACACCAAGACGTGAAGCAGCTGACCGAGGCGCTGCGAAGATCG 1798
Db 3616 GAATGAAGGCTGCCACACTAATGATGTGAAACAATTAACAGAGGAGGTACAAAAAATAG 3675
QY 1799 CCATGAGAGCATCTGTATCTGGGGCAAGACCCCAAGTTCCCGCTGCGCCATCCAGAAAG 1858
Db 3676 CCACAGAAAGCATAGTAATATGGGGAAGAACTCTTAATTTAAATTTACCATCAAAAAGG 3735
QY 1859 AGACCTGGGAGACCTGGTGGACGACTACTTGGCAGGCCACCTTGATCCCGAGTGGAGT 1918
Db 3736 AAACATGGGAGCATGGTGGACAGACTATTGGCAAGCCACCTGGATCTCTGAGTGGAGT 3795
QY 1919 TCGTGAACACCCCGCCCTTGGTGAAGCTGTGTGTACAGCTGGAGAGGAGGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTTCTTGTAGTGAAGTTATGTTACCACTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCCGAGACCTTCTACGTGGAGGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCG 2038

Db 3856 GAGCAGAAACTTICTATGTAGTGGCGAGCAATAGGAAACTAAATTAGGAAGAGCAG 3915
QY 2039 GCTACGTGACGCGCGCGGGCGGAGAGATCGTGGAGCCTGACCGAGACCAACACAGCAG 2098
Db 3916 GATATGTAACCTGACAGAGAGAGACAAAAGTTGTCCCCCTAACGAGACAAACAATCAGA 3975
QY 2099 AGACCGAGCTGCAAGGCCATCCAGCTGCGCCCTGCAAGGACAGCGGAGGTGAACATCG 2158
Db 3976 AGACTGAGTTACAAAGCAATTCATCTAGCTTTGCAAGGATTCGGGATTAGAAAGTAACATAG 4035
QY 2159 TGACCCGACAGCAGTACGCGCTGGGCATCTCCAGGGCCCGACCAAGCGGAGGAGCG 2218
Db 4036 TGACAGACTCACAATATGCTATGGAATCATCAAGCACAACACAGATAAGAGTGAATCAG 4095
QY 2219 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACTGTAGCTGGG 2278
Db 4096 AGTTAGTCACTCAAAATAATAGAGCAGTTAATAAAAAGGAAAAGTCTAGCTGGCATGG 4155
QY 2279 TGCCCGCCCAAGGGCATCGCGGCAACGAGGAGATCGACAAGCTGCTGAGCAAGGCA 2338
Db 4156 TACAGCACAACAAGGAATTTGGAGGAATGAACAAGTAGATGGTTGTCAGTCTCGAA 4215
QY 2339 TCCGCAAGCTGTGTTCTCTGGACGGCATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247
RESULT 9
US-09-309-572-15
; Sequence 15, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 9709
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-309-572-15
Query Match 45.6%; Score 1125.6; DB 4; Length 9709;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;
QY 14 TGCCCGGCGCCATGAGCCAGGCCACCA---GCGCCACATCTCTGTGTCAGCGCAGCACT 70
Db 1877 TGCTGAAGCAATGAGCCAGTAACTAACAAATCCAGCTACCATTAATGATACAGAAAGCAATT 1936
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCGGCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTGCAAGAGAGGCGCATAGCCA 1996
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGTGTTGGAAGTSCGGAAGGAGGCGCACCCAGA 190
Db 1997 AAAATTGCAAGGCGCCCTAGGAAAAGGGCTGTGGAATAATGTAAGAAAGGAGGACCCAAA 2056
QY 191 TGAAGACTGTCACCGAGCGCCAGGCCCAACTTCTTCGCGAGGACCTTGGCCTTCCCCCAGG 250
Db 2057 TGAAGATTGTTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCCAAC 2115
QY 251 GCAAGCGCCCGGAGTTCCCGAGCGAGCAACCGCCCAACAGCCACCCACCCAGCGGAGC 310
Db 2116 GGAAGCGCAGGGAATTTCTTCAGAGCAGACACAGACCCACCCACCCAGAGAGAGC 2175
QY 311 TCGAGTGGCGG-----CGACAACCCCGCGCAGCGAGCGCGCGCGCGAGCGCGGCA 364

Db 2176 TTTCAGGTTTGGGGAAGAGACACAACTCCCTCTCTCAGAAAGCAGGAGCCGATAGACAAGGAA 2235
QY 365 -----CCCTGAACCTTCCCCAGATCACCTCTGTGCAGCGCCCTGCTGGTGCATCAAGG 418
Db 2236 CTGTATCCCTTACCTTCCCTCAGATCACTCTTGTGGCAGGAGCCCTCGTCAAAAGAA 2295
QY 419 TGGCGGCGCAGATCAAGGAGGCCCTGCTGCAGACCGGCGCCGACACACCGCTGCTGGAGG 478
Db 2296 TAGGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAG 2355
QY 479 AGATGAGCTGCCCGGCAAGTGAAGCCCAAGATGATGGCGGCGATCGCGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAGATGGAACCAAAATGATAGGGGGAATTTGGAGGTTTTATCA 2415
QY 539 AGGTGGCGCATACCAACCATCTGATCGAGATCTGGCGCAAGGAAGGCCATCGGCACCG 598
Db 2416 AAGTAGGACAGTATGATCAGATACCTATAGAAATCTCGGGACATAAGAGTATAGGTACAG 2475
QY 599 TGCTGATCGGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGAOCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGGCA 718
Db 2536 GCACCTTAAATTTCCCATTTAGTCTTATGAGACTGTACCAAGTAAATTAAGCCAGGAA 2595
QY 719 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGGCCCAAGGTTAAACATGGCCATTGCACAGAGAAATAAAGCATTAGTAG 2655
QY 779 CCATCTGCAGAGATGGAGAGGAGGCAAGATCACCAAGATCGGCCCGCCGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGGAAGGAAATTTCAAAATTTGGCCTGAAATTCAT 2715
QY 839 ACAACACCCCGTGTCCCATCAGAGAGAGAGGAGCAGCACCAGTGGCGCAGCTGGTG 898
Db 2716 ACAATCTCCAGTATTGCCATTAAGAAAGAAAGACAGTACTAAATTTGGAGAAATTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAGGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
Db 2776 ATTTTCAGAGAACTTAATAAGAGACTCAAGATTCTGGGAAGTTCATTAGGAATACCAC 2835
QY 959 ACCCGCGCGCTGAGAGAGAGAGACGCTGACCGTCTGAGCTGGGCGACGCTTACT 1018
Db 2836 ATCTGCGAGGTTTAAACAGAGAAATACAGTAACAGTACTAAATTTGGAGAAATTAGTAG 2895
QY 1019 TCAGGCTGCCCTGGACGAGACTTCCCAAGTACACCGCTTCACCATCCCGACATCA 1078
Db 2896 TTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATCTGCAATTTACCATCTAGTATAA 2955
QY 1079 ACAAGGAGACCCCGGATCCGCTACCAAGTACAACGTGCTGCCCGAGGGCTGGAAGGCA 1138
Db 2956 ACAATGAGACACCAAGGATTAGATATCAGTACAATGTCTTCCACAGGGATGGAAGGAT 3015
QY 1139 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCTCGGACCCCTTCCGCGCCCGCAACC 1198
Db 3016 CACCAGCAATATCCAGTGTAGCATGACAAAATCTTAGAGCCCTTTAGAAAACAAATC 3075
QY 1199 CCGAGATCGTGATCTACCAATACATGAGACGCTGTAGTGGCGAGCGACCTGGAGATCG 1258
Db 3076 CAGACATAGTATCATCAATACATGATGATTTGTATGTAGGATCTGACTTAGAAATAG 3135
QY 1259 GCCAGCACCGCGCAAGATCGAGGAGCTGGCGAAGCAGCTCTCGCTGGGGCTTCACCA 1318
Db 3136 GGCAGCATAGAAAAAATAGAGCACTGAGCAACAATCTCTGAGGTGGGATTTACCA 3195
QY 1319 CCGCGCAGAGAACACACAGAGAGGCCCCCTTCCTTGTGGATGGCTTACAGCTGACG 1378
Db 3196 CACCAGCAAAAAACATCAGAAAGAACTCCATTCCTTTGGATGGGTATGAATCCATC 3255
QY 1379 CCGAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAGAGAGCTGGACCGCTGACG 1438

Db 3256 CTGATAAATGGACAGTACAGCCTATAGTCTGCCAGAAAGGACAGCTGGACTGTCAATG 3315
QY 1439 ACATCCAGAAGCTGTGGGCAAGCTGAATGGGCCAGCAGATCTACCCCGGCGATCAAGG 1498
Db 3316 ACATACAGAAATTAGTGGGAAATTAATTTGGSCAAGTCAGATTTATGCAAGGATTAAG 3375
QY 1499 TCGGCCAGCTGTGCAAGCTGCTCGCGGCGCCAAAGGCCCTGACCGACATCGTCCCTGA 1558
Db 3376 TAAGCAATTTATTAACCTTTAGGGGAACCAAGACACTAACAGAAAGTAGTACCCTAA 3435
QY 1559 CCGAGGAGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTCGCGAGCCCGCTGACG 1618
Db 3436 CAGAAGAAGCAGAGCTAGAACTGGCAGAAACACAGGAGATCTTAAAGAACCCTGACATG 3495
QY 1619 GCGTGTACTACGACCCCAAGCAGGACCTGGTGGCCGAGATCCAGAGAGAGGGCCAGCC 1678
Db 3496 GAGTGTATTTATGACCCATCAAAAGACTTAATAGCAAAATACAGAAAGCAGGSCAAGGCC 3555
QY 1679 AGTGACCTTACCAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1738
Db 3556 AATGACATATCAATTTATCAAGAGCCATTTAAANAATCTGAAACAGGAAATATGCAA 3615
QY 1739 AGATGGCAGCCGCCACACACAGCAGTGAAGCAGTGAACCGGAGCGCTGCAGAAAGATCG 1798
Db 3616 GAATGAAGGTTGCCACACTAATGATGTGAACAAATTAACAGAGCAGTACAAAAATAG 3675
QY 1799 CCATGGAGAGCATCGTGTCTGGGCAAGACCCCAAGTTCCGCTGCCCTCCAGAAAG 1858
Db 3676 CCACAGAAGCATAGTAAATATGGGAAGAGACTCTTAATTTAAATTTACCCATACAAAAG 3735
QY 1859 AGACCTGGGAGACCTTGTGGACCGCTACTGTCAGGCGCACCTTGGATCCCGAGTGGGACT 1918
Db 3736 AAACATGGAGCATGTTGGACAGAGTATTGCAAGCCACTTGGATTTCTGAGTGGGAT 3795
QY 1919 TCGTGAACACCCCGCTGTTGAGCTGTGTACAGCTGGAGAGAGGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTTACAGTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GGGCGGAGACCTTCTACGTGGAGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCCG 2038
Db 3856 GAGCAGAACTTTCTATGTAGATGGGCGAGCCATAGGGAACCTAAATTTAGGAAAAGCAG 3915
QY 2039 GCTAGTCAACGACCGGGCGCGCAGAGATCGTGAGCTGACCGAGACCCAGCCCAACAG 2098
Db 3916 GATATGTAACAGAGAGAGACAAAGTTGTCCCCCTAACGGACACACAAATCAGA 3975
QY 2099 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGACAGCGGAGCGAGGTGAACATCG 2158
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTGCGAGGATTCGGGATTAAGTAAACATAG 4035
QY 2159 TGACCGAGCGCAGTACGCCCTTGGGCATCATCCAGGCCCGCCGACAGAGCGGAGCG 2218
Db 4036 TGACAGACTCAATATGCTTGGGAATCATTTCAAGCACAACCATAGATAGATGAATCAG 4095
QY 2219 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGG 2278
Db 4096 AGTTAGTCAGTCAATAATAGAGCAGTTAATAAAAAAGAAAAAGTCTACCTGGCATGG 4155
QY 2279 TGCCCGCCCAAGGGCATCGCGCAAGCAGGACATCGCAAGCTGGTGAAGCAAGGCA 2338
Db 4156 TACCAGCACAAAAGGAATTTGGAGAAATGAACAAAGTATAGTGGTTGGTCAAGTGGAA 4215
QY 2339 TCCGCAAGGTGCTGTTCTCTGGACGCGATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTAGATGAATAGAT 4247

RESULT 10

us-08-935-312-13
; Sequence 13, Application US/08935312
; Patent No. 6207455
; GENERAL INFORMATION:
; APPLICANT: CHANG, Lung-Ji

;; TITLE OF INVENTION: LENTIVIRAL VECTORS
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NETMARK, P.L.L.C.
;; STREET: 624 Ninth Street, N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20001
;; COMPUTER READABLE FORM: disk
;; MEDIUM TYPE: Floppy
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA: US/08/935,312
;; APPLICATION NUMBER: US/08/935,312
;; FILING DATE: 22-SEP-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: COOPER, Iver P.
;; REGISTRATION NUMBER: 28,005
;; REFERENCE/DOCKET NUMBER: CHANG=112
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 12494 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: other nucleic acid
;; DESCRIPTION: /desc = "DNA"
;; US-08-935-312-13

Query Match 45.6%; Score 1125.6; DB 4; Length 12494;
Best Local Similarity 68.4%; Pred. No. 1.1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

Qy 14 TGGCCGAGGCATGAGCCAGCCACCA--CGGCCACATCTCTGATGAGCGCAGCAACT 70
Db 1548 TGGCTGAAGCAATGAGCCAGTAACAAATCCAGTACCATTAATGATACAGAAAGGCRAAT 1607

Qy 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGCGCCACATCGCCC 130
Db 1608 TTAGAACCAAGAAAGACTGTTAAGTGTTCATTTGGCAAGAGAGGGGCACATAGCCA 1667

Qy 131 GCAACTGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCCACAGGA 190
Db 1668 AAAATTGCGGGCGCCCTAGGAAAAAGGCTGTTGGAATGTGGAAGAGGAGGACACCAAA 1727

Qy 191 TGAAGGACTGACCGAGGCGCCAGCCCACTTCTCCGCGAGGACCTGGCTTCCCGCAGG 250
Db 1728 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTATGAGGAAGATCTGGCTTCCCAAG 1786

Qy 251 GCAAGGCGCGGAGTTCCCAAGCAGCAGAGAACCGCCCAAGCCCAAGCCCGCGAGC 310
Db 1787 GGAAGCCAGGGAATTTCTTTCAGAGCAGACAGGCCACAGCCCGCCACAGAGAGAGC 1846

Qy 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGGCGCGCGCGCGCGCGCGCGAGGCA 364
Db 1847 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAAAGCAGGAGCGCATAGACAAAGGA 1906

Qy 365 -----CCCTCAACTTCCCGCAGATCACCTGTGCGAGCGCCCTGGTGAGCATCAAGG 418
Db 1907 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCGTCAATTAAGA 1966

Qy 419 TGGCGGCGCAGATCAAGAGGCGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGTGGAGG 478
Db 1967 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACATATTAGAAAG 2026

Qy 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGCGGGATCGCGCGCTTCATCA 538
Db 3107 CAGAAGAGCAGAGCTAGAACTGGCAGAAACAGGAGGATTTCTAAAGAACCGGTACATG 3166

Db 2027 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 2086

Qy 539 AGGTGCCAGTAGCAGCACAGATCCTGATCGAGATCTCGGCAAGAGCCATCGGACCG 598
Db 2087 AAGTAGGACAGTAGATCAGATACATCTATGAAATCTCGGACATAAAGCTATAGGTACAG 2146

Qy 599 TGTGTATGCGGCCCGCCCGTGAACATCATCGGCCCGCAACATCTGACCCAGCTGGGCT 658
Db 2147 TATTAGTAGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2206

Qy 659 GCACCTGAACTTCCCGCATCAGCCCCATCGAGACCGTGGCGGTGAAGCTGAAGCCCGCA 718
Db 2207 GCACCTTAAATTTTCCCATTAGTCTTATGAGCTGTACCAAGTAAATTAAGCCAGGAA 2266

Qy 719 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2267 TGGATGGCCCAAAAGTTAAACAATGGCCATTTGACAGAGAGAAAAAATAAGCATTAGTAG 2326

Qy 779 CCATCTCGAGGAGATGGAGAAAGAGGCGCAAGATCACCAGATCGGCCCGCGAGAACCCCT 838
Db 2327 AATTTGTACAGAAATGGAAGAGGAGGAAAAATTTCAAAATTTGGCCCTGAAATCCAT 2386

Qy 839 ACAACACCCCGTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
Db 2387 ACAATACTCCAGTATTTGCCATAAAGAAAAAGACAGTACTTAAATGGAGAAAAATTAGTAG 2446

Qy 899 ACTTCGCGGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCGATCCGCC 958
Db 2447 ATTTGAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAAGTTCAATTAGAAATACCAC 2506

Qy 959 ACCCGCGCGGCTGAAGAAGAAAGAGCGGTGACCGTCTGGAGCTGGGCGCACCCCTACT 1018
Db 2507 ATCTGCGAGGTTTAAACACAGAAAAAATCAGTAACAGTACTGGATGTGGCGCATGATATT 2566

Qy 1019 TCACGCTGCCCTGGACGAGGACTTCGCCAAGTACACCGCCTTCACCATCCCGACATCA 1078
Db 2567 TTTGAGTTCCTTAGATAAAGACTTCAGGAAGTATCTGCACTTTACCATCTAGTATAA 2626

Qy 1079 ACAACGAGACCCCGGATCGCTACCACTACAAGCTGCTGCCCCAGCGCTGGAGGGCA 1138
Db 2627 ACAATGAGACACAGGATTAGATATCAGTACAAATGTCTTCCACAGGATGGAAGGAT 2686

Qy 1139 GCGCCAGATCTTCAGAGCAGCATGACCAAGATCTCTGAGCCCTTCGCGCGCCGCAACC 1198
Db 2687 CACGAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC 2746

Qy 1199 CCGAGATCGTATCTACAGTACATGGACGACCTGTACGTGGCGACGACCTGGAGATCG 1258
Db 2747 CAGACATAGTCTATCTCAATACATGGATGATTTGTATGATAGGATCTGACTTAGAATAG 2806

Qy 1259 GCCAGACCGCGCCAAAGATCGAGGAGCTGCGCAAGCACTGCTGCGCTGGGCTTCAACA 1318
Db 2807 GCGAGCATAGAACAAAAATAGGNACTGACACACATCTGTTGAGGTGGGATTTTACCA 2866

Qy 1319 CCCCGCAGACAGAACACACAGAGGAGCCCGCTTCTGATGGATGGCTTACAGCTGCAAC 1378
Db 2867 CACGAGCAAAAAACATCAGAAAAAACCCTCCATCTCTTTGGATGGGTATTGAACCTCCATC 2926

Qy 1379 CCGACAGTGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAAGC 1438
Db 2927 CTGATAAATGACAGTACAGCTATAGTGTGCCAAAGAGACAGCTGGACTGTCAATG 2986

Qy 1439 ACATCCAGAACTGTTGGGCAAGCTGAATGGGCGCAGCAGATCTTACCCCGCATCAAGG 1498
Db 2987 ACATACAGAAATAGTGGGAAAAATTTGAATTTGGCAAGTCAGATTTATGAGGGATTAAG 3046

Qy 1499 TGGCGCAGCTGTGAAGCTGTGCGCGCGGCGCAAGCCCTGACCCACATCTGTCGCCCTGA 1558
Db 3047 TAAGGCAATATGTAAGCTTCTTAGGGGAACCAAAAGCACTAACAGAAAGTAGTACCACATA 3106

Qy 1559 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGCCCGGCGACG 1618
Db 3107 CAGAAGAGCAGAGCTAGAACTGGCAGAAACAGGAGGATTTCTAAAGAACCGGTACATG 3166

```
QY 1619 GCGTGTACTACGACCCAGCAGGACCTGGTGGCGGAGATCCAGACGAGCGCCACGACC 1678
Db 3167 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCGGGAAGGCC 3226
QY 1679 AGTGACCTACCAGATCTACAGGAGCGCTTCAAGAAGCTGAAGACCGGCAAGTACGCCA 1738
Db 3227 AATGACATATCAAAATTTATCAAGGCCATTTAAATCTGAAAACAGGAAATATGCA 3286
QY 1739 AGATGCCGACCCGACACACACACGACGCTGAAGCAGCTGACCGGCGCTGAGAGATCG 1798
Db 3287 GAATGAAGGGTGGCCACACTAATGATGTGAACAAATTAACAGAGCGAGTACAAAAATAG 3346
QY 1799 CCATGGAGAGATCGTGTATCTGGGCGCAAGCCCAAGTTCCTGCGCTGCGCCATCAGAAAG 1858
Db 3347 CCACAGAAAGCATAGTAAATATGGGGAAGACTCTTAATTTAAATATCCCATACAAAGG 3406
QY 1859 AGACCTGGAGACCTGGTGGACCGACTACTGGCAGGCGACCTGGATCCCCAGTGGGAGT 1918
Db 3407 AAACATGGGAAGCATGGTGGACAGAGTATGGCAAGCCACCTGGATTCTGAGTGGGAGT 3466
QY 1919 TCGTGAACACCCGCCCTGGTGAAGCTGTGTACAGCTGAGAGGAGGAGCCCATCATCG 1978
Db 3467 TTGTCAATACCCCTCCCTTTAGTGAAGTTATGGTACCAAGTTAGAGAAAGAACCCATAATAG 3526
QY 1979 GCGCGGAGACCTTCTACGTGGACGCGCGCGCAACCGCGAGACCAAGATCGGCAAGGCG 2038
Db 3527 GAGCAGAAACTTTCTNTGTAGATGGGCGAGCAATAGGAAACTTAATAGGAAAGCAG 3586
QY 2039 GCTACGTGACCGACCGGGCGGCGGAGAGATCGTGAAGCTTACCGGAGACCAACACAGA 2098
Db 3587 GATATGTAAGTACAGAGGAGAGACAAAAGTTGTCCCGCTAAACGACACAAATCAGA 3646
QY 2099 AGACCGAGCTGCAGGCGCATCCAGTGGCCCTGCGAGGAGCGCGGAGGAGTGAACATCG 2158
Db 3647 AGACTGAGTTACAAGCAATTCATCTAGCTTTGAGGATTCGGGATTTAGAATTAACATAG 3706
QY 2159 TGACCGAGCAGCAGTAGCCCTGGGCGCATCTCCAGGCGCGCGGCAAGAGCGAGAGCG 2218
Db 3707 TGACAGACTCACAATATGATCTGGGATCAATTCAGCAGCAACAGATAGAGTGAATCAG 3766
QY 2219 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG 2278
Db 3767 AGTTAGTCAGTCAATATAGAGCAGTTAATTAAGGAGGAGGAGGAGTGTACCTGAGCTGG 3826
QY 2279 TGCCGCGCCACAGGCGCATCGCGGCAACGAGCAGATCGCAAGCTGTGAGCAGAGGCA 2338
Db 3827 TACCAGCACAAAGGATTTGGAGGAATGACAAAGTAGATGGTGGTTCAGTCTCTGAA 3886
QY 2339 TCCGCAAGGTGCTTCTCTGGACGGCATCGAT 2370
Db 3887 TCAGGAAGTACTATTTTATGATGGAATAGAT 3918
```

RESULT 11

```
US-08-848-760B-33
; Sequence 33, Application US/08848760B
; Patent No. 6248721
```

GENERAL INFORMATION:

```
; APPLICANT: Chang, Lung-Ji
; TITLE OF INVENTION: Animal Model For Evaluation Of Vaccines
```

```
; NUMBER OF SEQUENCES: 33
```

```
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
```

```
; STREET: 2421 N.W. 41st Street, Suite A-1
```

```
; CITY: Gainesville
```

```
; STATE: Florida
```

```
; COUNTRY: United States of America
```

```
; ZIP: 32606
```

```
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
```

```
; FILING APPLICATION NUMBER: US/08/848,760B
```

```
; FILING DATE: 25-Jan-2001
```

```
; CLASSIFICATION: <Unknown>
```

```
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: 08/838,702
```

```
; FILING DATE: 09-APR-1997
```

```
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: PACE, DORAN R.
```

```
; REGISTRATION NUMBER: 38,261
```

```
; REFERENCE/DOCKET NUMBER: CNG-100C1
```

```
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: (352) 375-8100
```

```
; TELEFAX: (352) 372-5800
```

```
; INFORMATION FOR SEQ ID NO: 33:
```

```
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 12494 base pairs
```

```
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
```

```
; TOPOLOGY: circular
```

```
; MOLECULE TYPE: other nucleic acid
```

```
; DESCRIPTION: /desc = "DNA"
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 33:
```

```
US-08-848-760B-33
```

```
Query Match 45.6%; Score 1125.6; DB 4; Length 12494;
Best Local Similarity 68.4%; Pred. No. 1,161,176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;
```

```
QY 14 TGGCCGAGGCGCATGAGCCAGGCCACCA---GCGCCAAACATCTGTGATGAGCGCAGCAACT 70
Db 1548 TGGCTGAAGCAATGAGCCAGCAAGTAACAATCCAGCTACCATATGATACAGAAAGCAATT 1607
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAAGTCGCGGAAGAGAGGCGCACATCGGCC 130
Db 1608 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTCGAAGAGAGGCGACATAGCCA 1667
QY 131 GCAACTGCCGCGCCCCCGCAAGAGAGGCTCTGGAAGTGGCGCAAGAGGCGCACACAGA 190
Db 1668 AAAATTGCGAGGCGCCCTAGGAAAGAGGCTCTTGGAAATGTGGAAGAGAGGACACCAAA 1727
QY 191 TGAAGGAGTGCACCGAGCGCCAGCCAACTTCTTCCGCGAGGACCTGCGCTTCCCGCAGG 250
Db 1728 TGAAGAGTTGATGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGCGCTTCCACAG 1786
QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGCAGAGAACCGGCCAACAGCCCGCCAGCGCGAGC 310
Db 1787 GGAAGGCGAGGGAATTTCTTCAGAGCAGACAGACCAACAGCCCGCCAGAGAGAGC 1846
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGGAGCGCGCGCGCGCGCGCGCGCA 364
Db 1847 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAAAGAGAGCGCGCATAGACAAGGAA 1906
QY 365 -----CCCTGAAGTTCGCCAGATCACTCTGTGCGAGCGCCCTGTGTGAGCATCAAG 418
Db 1907 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCTGTCAATAAAGA 1966
QY 419 TGGCGCGCGAGATCAAGAGGCGCTCTGTGGACACCGCGCGCGCGAGACACCTGTGTGAGG 478
Db 1967 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2026
QY 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGCGGCGCATCGCGCGCTTCATCA 538
Db 2027 AAATGAATTTGCCAGGAAGATGGAAACCAAAATATAGGGGGGAATTTGAGGTTTATCA 2086
QY 539 AGTGCGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 2087 AAGTAGACAGATGATGATCACTACTCATAGAAATCTCGGAGACATTAAGCTATAGTACAG 2146
QY 599 TGTGTATCGGCGCCACCCCGGTGAACATCATCGGCCCGCAACATCTGACCCAGCTGGGCT 658
Db 2147 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGTGACTCAGATTGGCT 2206
```

QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAAGCTGAAGCCCGGCA 718
DB 2207 GCACTTAAATTTCCCATAGTCTTATGAGACTGTACCACTAAATTAAGCCAGAA 2266
QY 719 TGGAGCCGCCAAGGTGAAGAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 2267 TGGAGCCGCCAAGGTGAAGAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 2326
QY 779 CCATCTGCGAGAGATGGAGAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCT 838
DB 2327 AAATTTGTACAGAAATGGAAAGAGAGGAAATTTCAAAATTTGGCCCTGAAATCCAT 2386
QY 839 ACAACACCCCGTGTGCGCATCAAGAAAGAGACAGCACCAAGTGGCGCAAGCTGGTG 898
DB 2387 ACATACTCCAGTATTTGGCATTAAGAAAGAGACAGTACTAATAGGAGAAATTAGTAG 2446
QY 899 ACTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGACAGTGGGATCCCC 958
DB 2447 ATTTACAGAACTTAATAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCC 2506
QY 959 ACCCGCGCGCTGAAGAGAGAGAGCGGTGACCGTGTGAGCTGGGCGAGCCCTACT 1018
DB 2507 ATCCTGCAGGGTTTAAACAGAGAAATACAGTAAAGTACTGGATGTGGCGATGCAAT 2566
QY 1019 TCAGCGTGGCCCTGGAGGAGCTTCGCAAGTACACCGCTTTCACCATCCCCAGCATCA 1078
DB 2567 TTTCACTTCCCTTAGATAAGACTTCAGGAGTATACATTTACCATACCTAGTATAA 2626
QY 1079 ACAACAGAGCCCCGGATCCCGTACAGTACAACTGTGTCGCCCGAGGGTGGAGGGCA 1138
DB 2627 ACAATGAGACCCAGGATTTAGATATCAGTACAAATGTGCTTCCAGGGATGGAAGGAT 2686
QY 1139 GCGCCAGCATCTTCAGAGAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCGCAACC 1198
DB 2687 CACAGCAATATTTCCAGTGTAGCATGACAAATCTTAGAGCCCTTTAGAAAACAAATC 2746
QY 1199 CCAGATCGTGTATCTACAGTACATGAGACCTGTACCTGGCGAGCGACTGGAGATCG 1258
DB 2747 CAGCATAGTCACTATCAATACATGATGATTTGTATGTAGTACGTGACTTAGAATA 2806
QY 1259 GCAGACCCGCGCAGATCGAGAGCTGGCGAAGCACTGTGCGTGGGGCTTACCA 1318
DB 2807 GCGAGATAGAAACAAATAGAGAACTGAGACAACTCTGTTGAGTGGGGATTTACCA 2866
QY 1319 CCGCCGACAGAGACACAGAGGAGCCCTTCTGTGGATGGCTAGAGCTGCACC 1378
DB 2867 CACGACAAACAAATCAGAAAGAACCTTCCATTTCTTGGATGGGTATGAACTCCATC 2926
QY 1379 CCAGCAAGTGGACCGTGCAGCCCATCGAGTGCCTCGAGAGAGAGAGTGGACCGTGAACG 1438
DB 2927 CTGATAAATGGACAGTACAGCTTATAGTGTGCGCAGAAAGGACAGCTGGACGTCAATG 2986
QY 1439 ACATCCAGAGCTGTGGGCAAGCTGACTGGCCAGCCAGATCTACCCCGGATCAAGG 1498
DB 2987 ACATACAGAAATTTAGTGGGAAATTAATTTGGCAAGTCAAGTTTATGAGGGATTAAG 3046
QY 1499 TCGCCAGCTGTGCAAGCTGCTGGCGGCGCCAGGCGCTGACCGACATCGTGCCTGA 1558
DB 3047 TAGGCAATTTATGAACTTTTAGGGAAACCAAGCACTAACAGAGTAGTACCATAA 3106
QY 1559 CCGAGAGCGCGAGCTGGAGCTGGCGAGAACCCGAGATCTTGGCGGAGCCCGTGCACG 1618
DB 3107 CAGAAGAAGCAGAGCTAGAACTGGCAGAAACAGGGAGATTTCTAAAGAAACCGGTACATG 3166
QY 1619 GCGTGTACTACGACCCAGCAAGCACTGTGCGCGAGATCCAGAGCAGGCGCCAGCACC 1678
DB 3167 GAGTGTATTTAGCCCATCAATCAAAAGACTTAATAGCAAAATACAGAGCGGGCAAGGCC 3226
QY 1679 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGACCTTGAAGACCGGCAAGTACGCCA 1738
DB 3227 AATGGACATATCAAAATTTATCAGAGGCCATTTAAATCTGAAACAGGAAATATGCA 3286

QY 1739 AGATGCGCACCGCCACACACACAGCAGCTGAAGCAGCTGACCGAGGCGGTGCGAAGATCG 1798
DB 3287 GAATGAAGGCTGCCACACTAATATGATGTAAACAATTAACAGAGGCGGTACAAAAATAG 3346
QY 1799 CCATGAGAGAGCTGTGATCTGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1858
DB 3347 CCACAGAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAAATTTACCCATACAAAG 3406
QY 1859 AGACCTGGGAGACCTGGTGGACCGACTACTGCGAGCGCACTTGATCTCCCGAGTGGAGT 1918
DB 3407 AAACATGGGAAGCATGGTGGACAGAGTATTTGGCAAGCCACTGGATCTCTGAGTGGAGT 3466
QY 1919 TCGTAACACCCCGCCCTGGTGAAGCTGTGTACCAGCTGGAGAGGAGGCCCATCATCG 1978
DB 3467 TTGTCAATACCCCTCTCTAGTGAAGTTATGTTACCACTTGTAGAGAAAGAACCCATAATAG 3526
QY 1979 GCGCCAGAGCTTCTAGTGGACGCGGCCCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 2038
DB 3527 GAGCAGAACTTTCTATGTAGATGGGCGAGCCAATAGGGAACCTAATTTAGGAAAGCAG 3586
QY 2039 GCTACCTGACCGACCGCGCGCAGAAAGATCTGTAGCCCTGACCGAGACCCACCAACCAGA 2098
DB 3587 GATATGTAAGTACAGAGGAGACAAAAGTTGTCCCTTAACGGACACAAACATCAGA 3646
QY 2099 AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACACGCGGACGAGGTGAACATCG 2158
DB 3647 AGACTGAGTTACAAGCAATTCATCTAGCTTTGCAGGATTCGGGATTTAGAAAGTAAACATAG 3706
QY 2159 TGACCCACAGCAGTACGCTGGCATCATCCAGCCCGCCAGCCGACAGAGCGAGAGCG 2218
DB 3707 TGACAGACTCAACAATATGCAATGGGAATCATCAAGCACAACCAAGATTAAGAGTGAATCAG 3766
QY 2219 AGCTGTGTAACAGAGTATCATCGACGCTGATCAAGAGAGAGAGGTGACTGAGCTGGG 2278
DB 3767 AGTGTAGTCAAAATATAGACAGTATTAATAAAGGAAAGTCTACCTGCGCATGGG 3826
QY 2279 TGCCCGCCCAAGAGGCGATCGCGGCAACGAGAGATCGACAGCTGCTGAGCAAGGGCA 2338
DB 3827 TACCAGCACACAAGGAATTTGGAGGAATTAAGCAAGTAGATGGGTTGTCAGTCTGGAA 3886
QY 2339 TCCGCAAGGTGCTTCTCTGCGCGCATCGAT 2370
DB 3887 TCAGGAAGTACTATTTTAGTGGAAATAGAT 3918

RESULT 12

US-08-646-538-35
; Sequence 35, Application US/08646538
; Patent No. 6027881
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE: No. 6027881 yet assigned
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: -
LOCATION: 1..1581
OTHER INFORMATION: /note= "pNlnSG11"
US-08-646-538-35

Query Match 45.6%; Score 1125.6; DB 3; Length 1581;
Best Local Similarity 68.4%; Pred. No. 1..1e-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;
QY 14 TGGCCGAGCCATGAGCCAGGCCACCA--GCGCCAACATCCTGATGCGAGCGCAGCAACT 70
DB 1877 TGGCTGGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTCACTGAGCGGCAAGAGGGCCACATCGGCC 130
DB 1937 TTAGGAACCAAGAAAGACTGTGAAGTGTTCGAATTTGGCAAGAAAGGGCCACATAGCCA 1996
QY 131 GCAACTGGCGGCCCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACCCAGA 190
DB 1997 AAATTTGAGGGCCCCCTAGGAAAAGGGCTGTGGAATGTGGAAGAGGAGACACAAA 2056
QY 191 TGAAGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGAGACTGCGCTTCCCCCAGG 250
DB 2057 TGAAGATTGACTGAGAGACAGGGCTAA--TTTTTTAGGGAAGACTGTGGCTTCCCCACAAG 2115
QY 251 GCAAGGCCGGGAGPTCCCGAGGAGAGAACCCGCGCAAGAGCCCCACCCAGCCGAGC 310
DB 2116 GGAAGCCAGGGAATTTCTTCAGAGCAGACGAGGCAACAGCCCAAGCCCCACCCAGAGAGC 2175
QY 311 TGCAGGTGCGGG-----CGAACCCCGCAGGAGCGCGCGCCGAGCGCCAGGCGCA 364
DB 2176 TTCAGGTTGGGAAGAGACACAACTCCCTCTCAGAGCAGGAGCGCGATAGACAGGNA 2235
QY 365 -----CCCTGAACCTTCCCGAGATCACCTGTGCGAGCGCCCTGTGTGAGCATCAAGG 418
DB 2236 CTGTATCCTTTAGCTTCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAGA 2295
QY 419 TGGCGGCGCAGATCAAGAGGCGCTGCTGGACACCGCGCGCCGACACACCCGTGTGGAGG 478
DB 2296 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATCGGGGCAATCGCGGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGAATTTGGAGTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACAGATCTCTGATGCGAGATGTCGGCAAGAGGCCATFCGCAACG 598
DB 2416 AAGTAGGACAGTATGATCAGATCACTATAGAAATCTCGGGACATAAAGCTATAGGTACAG 2475
QY 599 TGCTGATCGGCCCCACCCCGTGRACATCATCGGCCGCAACATGCTGACCCAGCTGGGCT 658
DB 2476 TATTAGTAGGACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCCCTGAACCTTCCCATCATCGCCCATCTCAGACCGCTGCCGTGAAGCTGAAGCCCGCA 718
DB 2536 GCACCTTAATTTCCCATATTAGTCTATTGAGACTGTACCGAGTAAATTAAGCCAGGAA 2595
QY 719 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCCCTGACCGAGGAGAAAGATCAAGGCCCTGACCG 778

DB 2596 TGGATGGCCCAAAAGTTAAACAATGSCCATTTGACAGAAGAAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGGGAAGATCACCAGATCGGCCCGGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGGAAGGAAAGAAAATTTCAAAAATTTGGCCCTGAAAATCCAT 2715
QY 839 ACAACACCCCGCTGTTTCGCCATCAAGAAGAAGGACAGCACCAGTGCSCCAAGCTGCTGG 898
DB 2716 ACAATCTCCAGTATTGGCCATAAAGAAAAGACAGTACTAATAGGAAAAATTTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGAGTGCAGCTGGGCGATCCCC 958
DB 2776 ATTTTCAGAGAATTAATAAGAGAACTCAAGATTTCTTGGGAAGTTCAATTAGGAATACCAC 2835
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGAGCGTGACCGTGTGGACGTGGGCGGAGCGCTACT 1018
DB 2836 ATCTTCGAGGGTTAAACACAGAAAAATCACTAAGTACTGGATGTGGCGATGCAATATT 2895
QY 1019 TCAGCTGTCCTGGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCCCAGCATCA 1078
DB 2896 TTTTCAGTTCCCTTAGATAAAGACTTTCAGGAAGTATCTGCAATTTACCATACCTAGTATA 2955
QY 1079 ACAACGAGACCCCGGCAATCCGCTACCACTACAACTGCTGCCCGCAGGGCTGGAAGGCA 1138
DB 2956 ACAATGAGACACACAGGATTAGATATCAGTACAACTGCTTCCACAGGATGAAAGGAT 3015
QY 1139 GCCCGCAGATCTTCAGAGCAGCATGACCAAGATCTGGAGCCCTTCCGCGGCCGCAACC 1198
DB 3016 CACCAAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCTTTTAGAAAAACAATC 3075
QY 1199 CCGAGATCTGTATCTACCACTGACGAGCTGTACCTGGCGAGCTGTACCTGGAGCTGGAGATG 1258
DB 3076 CAGACATAGTCACTATCAATCAATGATGATTTGTATGATGATCTGACTTAGAATAG 3135
QY 1259 GCCAGCACCGCCCAAGATCGAGAGCTGCGCAAGCACCTGCTGCGCTGGGCTTCCACCA 1318
DB 3136 GGCAGCATAGAACAAAAATAGAGAACTGAGCAACATCTGTTGAGTGGGATTTACCA 3195
QY 1319 CCCCAGACAAAGAACACCAAGAGGAGGCCCTTCTCTGTGGATGGCTAGGAGCTGCACC 1378
DB 3196 CACCAAGAAAAAATCACTAGAAAGAACCTTCCATTTCTGATGGTGTATGAACTCCATC 3255
QY 1379 CCGACAAGTGGAGCTGTCAGCCCATCGAGCTGCCCGAGAGAGAGCTGACCGTGAACG 1438
DB 3256 CTGATTAATGGACAGTACAGCCTATAGTCTGCCAGAAAGACAGCTGACTGTCAATG 3315
QY 1439 ACATCCAGAAGCTGTGGCAAGCTGAATGGGCCAGCCAGCATCTACCCCGGCATCAAG 1498
DB 3316 ACATACAGAAATTAGTGGGAAAAATTGAATTGGCAAGTCAGATTTATGCAGGATTTAAAG 3375
QY 1499 TCGCGCAGCTGTGCAAGCTGTGCGGGGGCCCAAGCCCTGACCGACATCTGCGCCCTGA 1558
DB 3376 TAAGCAATATTGTAACCTTCTTAGGGGAAACCAAGCACTAACAGAAAGTAGTACCATAA 3435
QY 1559 CCGAGAGGCGCAGCTGGAGCTGGCGGAGAACCGGAGATCTCTCGCGAGCGCGTGCAGC 1618
DB 3436 CAGAGAAGCAGAGCTAGACTGGCAGAAACAGGAGATTTAAAGAACCCGTTACATG 3495
QY 1619 GCGTGTACTACGACCCCAAGAGGACCTGGTGGCGGAGATCCAGAGCAGGCGCCAGCC 1678
DB 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAAGAGGGCAAGGCC 3555
QY 1679 AGTGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTAGGCCA 1738
DB 3556 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAAATATGCAA 3615
QY 1739 AGATCGCAGCCCGCACCAACAGCAGTGAACAGCTGACCGAGCGCTGCGAGAGATCG 1798
DB 3616 GAATGAAGGTCGCCACACTATGATGTGAACAATTAACAGAGGCGAGTACAAAAAATAG 3675
QY 1799 CCATGGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCTTCCCATCCAGAGG 1858

Db 3676 CCACAGAAAGCATAGTAAATATGGGAAAGACTCTAAATTTAAATTACCCATACAAAAGG 3735
QY 1859 AGACCTGGAGACCTGCTGGACCACTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1918
Db 3736 AAACATGGGAGCATGCTGGACAGAGTATTTGGCAAGCCACCTGGATCTCTGAGTGGGAGT 3795
QY 1919 TCGTGAACACCCCCCTGGTGAAGCTGGTACCAGCTGCGAGAGAGAGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTTACCAAGTTAGAGAAAGAACCCATAATAG 3855
QY 1979 GCGCCGAGACCTTCTAGCTGGACGCGCCGCCCAACCGGAGACCAAGATCGCAGGCGG 2038
Db 3856 GAGCAGAAACTTTCTATGTAGATGGGCGCAATAGGGAACTAAATTTAGGAAAGCAG 3915
QY 2039 GCTACGTGACCGAGCGGGCGGCGAGAGATCGTGAGCTGACCGAGACCAACCAACAGA 2098
Db 3916 GATATGTAACAGAGGAGAGACAAAGATTGTCCTTAACGGACACACAATCAGA 3975
QY 2099 AGACCGAGCTGACGAGCATCATGCTGGCCCTGCGAGGACAGCGGAGAGGTGAACATCG 2158
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTCAGAGGATTCGGGATTAGAAAGTAAACATAG 4035
QY 2159 TGACCGACAGCAGTACGCTGGSCATCATCCAGGCGCCGCGACAGAGCGAGAGCG 2218
Db 4036 TGACAGACTCACATATGTCATTTGGGAATCATCAAGCAACACCAATAGAGTGAATCAG 4095
QY 2219 AGCTGTGAACAGCATCATGAGCAGCTGTATCAAGAGGAGAGGTGTACCTGAGCTGGG 2278
Db 4096 AGTTAGTCAGTCAATAATAGAGCAGTTAATAAAGAGAAAGATCTACCTGGCATGG 4155
QY 2279 TGCCCGCCACAGAGGATCGGCGGCGACAGAGAGATGACAGCTGCTGAGCAAGGCA 2338
Db 4156 TACCAGACACAAAGGAATTTGGAGGAATGAACAGTAGATGGTGGTCTGCTGGAA 4215
QY 2339 TCCCAAGGTCTGCTCTCTGAGCGCATCGAT 2370
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 13
US-09-503-222-35
; Sequence 35: Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/503,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..15581
; OTHER INFORMATION: /note= "pNlnSG11"
US-09-503-222-35

Query Match 45.6%; Score 1125.6; DB 4; Length 15581;
Best Local Similarity 68.4%; Pred. No. 1.le-176;
Matches 1622; Conservative 0; Mismatches 734; Indels 16; Gaps 4;

QY 14 TGGCCGAGGCGCATGAGCCAGGCCACCA---GCGCAACATCTCTGATGACGCGCAGCAACT 70
Db 1877 TGGCTGAGCAATGAGCCCAAGTAACAATCCAGTACCATANTGATACAGAAAGCAATT 1936
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAAGTTCGCGCAAGGAGGCGCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGAGACTGTTAAGTGTTCATTTGTCGCAAGAGGCGCACATAGCCA 1996
QY 131 GCAACTGCCGCGCCCGCCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCACCA 190
Db 1997 AAAATTGCAAGGCGCCCTAGGAAAAGAGGCTGTTGGAATTTGGAAGAGGAGGACCAAA 2056
QY 191 TGAAGGACTGCAAGGCGCGCAGCGCAACTTCTTCCGAGGAGCCTGGCTTCCCGCAGG 250
Db 2057 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGGCTTCCCAAG 2115
QY 251 GCAAGCGCGGAGTTCCCGCAGGAGCAACCGCGCAAGAGCCCAAGCCCGCAGCGCAGC 310
Db 2116 GGAAGCGCAGGGAATTTCTTCAGAGCAGACCAAGCTCCCTCTCAGAAAGCAGGAGCGATAGACAAGAA 2175
QY 311 TCGAGTGGCGG-----CGACAACCCCGCAGGAGGCGCGCGCGAGCGCCAGGCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAACACTCCCTCTCAGAAAGCAGGAGCGATAGACAAGAA 2235
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGAGCATCAAG 418
Db 2236 CTGTATCTTTCCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAGA 2295
QY 419 TGGCGCGCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGAGCACCGTGTGGAGG 478
Db 2296 TAGGGGGCAATTTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCCTGCGCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGGGCTTCATCA 538
Db 2356 AAATGAATTTTGGCAGGAAGATGGAACCAAAATGATAGGGGAAATTTGGAGGTTTATCA 2415
QY 539 AGTGCGCGCAGTACGACCGAGATCCTGATCGAGATCTCGGCAAGAGCCCATCGGCACCG 598
Db 2416 AAGTAGCAGATGATGATCAGATCACTCATAGAAATCTCGGACATAAAGCTATAGGTACAG 2475
QY 599 TGTGATCGGCGCCACCGCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGACCTACACCTGTCAACATAATTTGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGCGTGGCGCGTGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTTCCCATTTAGTCTTATGAGACTGTACCAAGTAAATTAAGCCAGAA 2595
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCGCCCTGACCGAGGAGAGATCAAGCGCTGACCG 778
Db 2596 TGGATGCGCCCAAGGTAAACATAGGCCATTGACAGAGAAAGAAATTAAGCAATTAGTAG 2655

Qy 779 CCATCTCGAGGAGATGGAGAGGAGGACCAAGATCCACCAAGATGGCCCGCCGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGAAAGGAGGAAATTTCAAAAATTTGGGCTTGAATAATCCAT 2715
Qy 839 ACAACACCCCGTGTTCGCCATCAAGAGAGAGACAGCACCAAGTGGCGCAGCTGGTGG 898
Db 2716 ACAATACTTCCAGTATTTGCCATAAGAAAGAAAGACAGTACTAAATGGAGAGAAATAGTAG 2775
Qy 899 ACTTCGGGAGCTGAACAAGCGCCACAGGACTTCTGGGAGGTGCAGCTGGGCGATCCCCC 958
Db 2776 ATTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAGTTCAATAGGAATACCAC 2835
Qy 959 ACCCGCGCGGCTTGAAGAAGAGAGCGGTGACCGTGTGGACGTGGCGACGCTACT 1018
Db 2836 ATCTGCGAGGTTTAAACAGAGAAAAATCAGTAACAGTACTGGATCTGGGCGATGATATT 2895
Qy 1019 TCAGGTGCCCTGGACGAGACTTCCCAAGTACACCGCTTCCACATCCCCAGCATCA 1078
Db 2896 TTTGAGTTTCCCTTATAGATAAGACTTCAGGAAGTATCTGCAATTTACCATACCTAGTATAA 2955
Qy 1079 ACAAGAGACCCCGGATCCGCTACCAAGTACAAGCTGCTGCCCGAGGCTGGGAAGGCA 1138
Db 2956 ACATGAGACACAGGATATAGATATCAGTACAATGTCTCCACAGGATGGAAAGAT 3015
Qy 1139 GCGGAGCATCTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCGCAACC 1198
Db 3016 CACGAGCAATATCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAATC 3075
Qy 1199 CCGAGATGCTATCACCAGTACATGGAGACCTGTAGCTGGGACGACGACCTGGAGTCG 1258
Db 3076 CAGACATAGTCATATCAATACATGGATGATTTGTATGTAGGATCTGACTTGAATAATAG 3135
Qy 1259 CCGAGACCGCGCAAGATCGAGAGCTGCGCAAGACCTGCTGCGTGGGCTTCACCA 1318
Db 3136 GCGAGCATAGACAAAATAGAGNACTGACACACATCTCTTGAGGTGGGATTTACCA 3195
Qy 1319 CCGCGACAGAACACACAGAGGAGGCCCCCTTCTGTGGATGGGCTACGAGCTGCACC 1378
Db 3196 CACGACAGAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGTTTATGAACCTCCATC 3255
Qy 1379 CCGACAACTGACCGTGCAGCCCATCGAGCTGCGGAGAGAGAGCTGGACCTGGAACG 1438
Db 3256 CTGATAATGACAGTACAGCTTATGTGTCGCAAGAAAGAGACGCTGGACTGTCAATG 3315
Qy 1439 ACATCCAGAACTGTGGGCAAGCTGAATCGGCGAGCAGATCTACCCCGCATCAAGG 1498
Db 3316 ACATACAGAAATTAGTGGGAAATTAATTTGGCAAGTCAATTTACGAGGATTAAG 3375
Qy 1499 TCGCCAGCTGTGAAGCTGTGCGCGCGCCCAAGGCCCTGACCGACATCGTGCCCTGA 1558
Db 3376 TAAGCAATTTATGTAACCTTCTTAGGGGACCAAGACACTAACAGAAGTAGTACCCTAA 3435
Qy 1559 CCGAGGAGCCGAGCTGGAGCTGCGGAGAACCGGAGATCTCTGGCGAGCCGCTGCACG 1618
Db 3436 CAGAAGAGCAGAGCTGAGACTGCGCAAGAAACAGGGAGATTCTAAAAGAACCGGTACATG 3495
Qy 1619 CGGTGTACTACGCCACCCAGCAGCCTGTGTGGCGGAGATCCAGAAGCAGGCGCACGACC 1678
Db 3496 CAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 3555
Qy 1679 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1738
Db 3556 AATGGACATATCAAAATTTATCAAGGCACTTTAAAATCTGAAACAGGAAATATGCA 3615
Qy 1739 AGATGCGCACCGCCACACCAACGACGTGAAGAGCTGACCGCCGCTGGAGAGATCG 1798
Db 3616 GAATGAAGGGTGGCCCACTAATGATGTGAACCAATTAACAGAGGCGAGTACAAAAATAG 3675
Qy 1799 CCATGGAGAGCATCGTGTATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAGG 1858
Db 3676 CCACAGAAAGCATAGTATATGGGGAAGACTCTCTAAATTTAAATACCCATACAAAGG 3735
Qy 1859 AGACCTGGGAGACCTGTGTGGACCGACTACTGGCAGGCGCACTGGATCCCCAGTGGGAGT 1918

Db 3736 AAACATGGGAAGCATGTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGAGT 3795
Qy 1919 TCGTGAACACCCCGCTTGTGTGAAGCTGTGTACAGCTGTGAGAGAGAGCCCATCATCG 1978
Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTGTACAGTTAGAGAAAGAACCCATATAG 3855
Qy 1979 GCGCGAGACCTTCTACCTGGACGCGCCCAACCGGAGACCAAGATCGGCAAGGCG 2038
Db 3856 GAGCAGAACTTTCTATGTAGTGGGAGCCCAATAGGAAACTTAATTAGGAAAGCAG 3915
Qy 2039 GCTAGTGAACGACCGGGCCGAGAGATGCTGTAGGCTTACCAGAGACCAACCAAGCA 2098
Db 3916 GATATGTAACAGAGAGGAGACAAAAAGTTGTCCCCCTAACGAGACACAAACATCAGA 3975
Qy 2099 AGACCGAGCTGCGAGGCGCATCCAGCTGGCCCTGCGAGACAGCGGAGGAGTGAACATCG 2158
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTGACAGGATTCGGGATTAAGAGTAAACATAG 4035
Qy 2159 TGACCGACAGCAGTACGCTGGGCTATCCAGGCGCAGCCGACAGAGAGGAGAGCG 2218
Db 4036 TGACAGACTCACAATATGCAATTTGGGAATCATTCAGACCAACCAAGATGAATCAG 4095
Qy 2219 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAAGAGAGAGTGTACCTGAGCTGG 2278
Db 4096 AGTTAGTCAGTCAAAATAATAGAGCAGTTAAATAAAGAGGAAAGTCTACCTGGCATGG 4155
Qy 2279 TGCCCGCCCAAGGCGATCGCGCGCAGGACGACATGCAAGCTGGTGAAGAGGCA 2338
Db 4156 TACCAGCACACAAAGAAATGGAGGAAATGAACAAGTAGAGTGGTGGTGTGCTGCTGGAA 4215
Qy 2339 TCCGCAAGGTGCTGTCTCTGGACGCGCATCGAT 2370
Db 4216 TCAGGAAAGTACTATTTTAGATGAATAGAT 4247

RESULT 14
US-08-944-449-7
; Sequence 7, Application US/089444449
; Patent No. 5985613
; GENERAL INFORMATION:
; APPLICANT: KURTH, REINHARD
; APPLICANT: BAIER, MICHAEL
; APPLICANT: METZNER, KARIN
; APPLICANT: WERNER, ALBRECHT
; TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
; TITLE OF INVENTION: lymphokine (ISL)" to inhibit the replication of
; FILE OF INVENTION: viruses, particularly of retroviruses
; FILE REFERENCE: 8341-7065
; CURRENT APPLICATION NUMBER: US/08/944,449
; CURRENT FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: EP 95113013.2
; EARLIER FILING DATE: 1995-08-18
; EARLIER APPLICATION NUMBER: DE 195 13 152.5
; EARLIER FILING DATE: 1995-04-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 9737
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-08-944-449-7

Query Match 45.2%; Score 1116; DB 2: Length 9737;
Best Local Similarity 68.1%; Pred. No. 4.le-175;
Matches 1616; Conservative 0; Mismatches 740; Indels 16; Gaps 4;
Qy 14 TGGCGAGCGCATGAGCCAGGCCACCA--GGCCCAACATCTGATGCGCCGACCAACT 70
Db 1884 TGGCTGAAGCCATGAGCCAGGTAACAAATCCAGCTAAACATATGATGAGAGGCAATT 1943
Qy 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTCAACTGGCGCAAGGAGGCGCCACATCGCCC 130

Db 1944 TTAGGAACCAAAAGAAAGACTGTTAAAGTGTTCATTTGTGGCAAGAGGGCCACATAGCCA 2003
QY 131 GCAACTGCCGCGCCCGCAAGAGGGCTGCTGGAAGTCGGCAAGAGGGCCACACAGA 190
Db 2004 AAAATTGCAGGGCCCTAGGAAAAGGGCTGTTGGAGATGTGGAAGGAGACACCAA 2063
QY 191 TGAAGACTGCACCGAGCGCCAGCCAACTTCTTCGCGAGGAGACCTGGCCCTTCCCCOAGG 250
Db 2064 TGAAGATTGCATGAGACAGAGCTAA-FTTTTATAGGAAGATCTGGCCCTTCTACAAG 2122
QY 251 GCAAGCCCGCGAGTTCGCCAGCGAGCAGAACCGCCGACACGCCCCACCGCGGAGC 310
Db 2123 GGAAGGCCAGGAAATTTTCTTCAGAGCAGACACAGAGCCACAGCCCGCCAGAGAGC 2182
QY 311 TGCAGGTGCGCGCG-ACAAACCCCGCAGCGAGCGCGCGCGCGAGCGCGAGGCA 364
Db 2183 TTCAGGTTTGGGAGGAGAAAACAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAA 2242
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTCTGGTGAGCATCAAGG 418
Db 2243 CTGTATCCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAATAAGGA 2302
QY 419 TGGCGGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGAGCACACCGTCTGGAGG 478
Db 2303 TAGGGGGCAACTAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGTATTAGAAG 2362
QY 479 AGATGAGCCTTGCCCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 538
Db 2363 AAATGAATTTGCCAGGAAATGGAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 2422
QY 539 AGTGGCCAGTACGACCGAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 2423 AAGTAAGACAGTAGATCACTGTAGAAATCTGTGACATAAAGCTATAGGTACAG 2482
QY 599 TGCTGTATGSCCCACCCCGTCAACATCATCGCGCGCAACATGCTGACCCAGCTGGCT 658
Db 2483 TATTAGTAGACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGTT 2542
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCOCGTGAAGCTGAAGCCCGGCA 718
Db 2543 GTACTTTAAATTTCCCATTTAGCTTATTGAAACTGTACCAGTAAAAATTAAGCCAGGA 2602
QY 719 TGACGGCCCCAAGTGAAGCACTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACC 778
Db 2603 TGGATGGCCCAAGATTAAGCAATGGCCATTTGACAGAAGAAAAATAAAGCATTAGTAG 2662
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 2663 AGATGTACAGAAATGGAAGAAAGAGGGAATTTCAAAAAATTTGGCCCTGAAATCCAT 2722
QY 839 ACAACACCCCGCTGTTCCGCATCAAGAAGAAAGACAGCAACCAAGTGGCGCAAGCTGGT 898
Db 2723 ACAATACTCCAGTATTGCTTATAAGAAAAAAGACAGTACTAAATGGAGAAAACTAGTAG 2782
QY 899 ACTTCGCGAGCTGAACACGCGACCCAGACACTCTGGGAGGTGACGCTGGGATCCCC 958
Db 2783 ATTTCAGAGAACTTAATAAAGAACTCAAGACTTCTGGGAAGTTCAGTTAGGAATACCAC 2842
QY 959 ACCCCGCGGCTGAAGAAGAAAGAGCGTGACCGTGTCTGGACGTGGGCGACCGCTACT 1018
Db 2843 ACCCGCAGGTTAAAAAGAAAAATCAGTAACAGTATTGGATGTGGGTGATGCATCT 2902
QY 1019 TCAGCGTGCCCTTGAGCGAGGACTTCGCAAGTACACCGCTTACCATCCCCAGCATCA 1078
Db 2903 TTTCACTTCCCTTAGATAAAGACTTTAGAAAGTATACTGCATTTACCATACCTAGTATAA 2962
QY 1079 ACAACGAGCCCCCGCATCCGCTACAGTACAACTGTGTCGCCCGAGGGCTGGAAGGGCA 1138
Db 2963 ACAATGAGACACACGGGATTTAGATATCAGTACAAATGTGTGTCACACAGGGATGGAAGGAT 3022
QY 1139 GCCCCAGCATCTTCCAGAGCAGATGACCAAGATCTCTGGAGCGCTTCCCGCGCCGAACC 1198
Db 3023 CACCAGCAATATTTCCAAAGTAGCATGACAAAAATCTTTAGAGCCCTTTTAGAAAAAGAAATC

QY 1199 CCAGAGATCGTGATCTACAGTACATGACGACACCTGTACGTGGCAGCGACCTGGAGATCG 1258
Db 3083 CAGACATAGTTATCTATCAATACATGATGATTGTATGTAGGATCTGCTAGAAATAG 3142
QY 1259 GCAGACACCGCGCAAGAGCTGCGAAGCACTGCTGCGCTGGGGCTTACCA 1318
Db 3143 GGCAGCATAGAACAAAATAGAGAACTGAGACAGCATCTGTTGAGGTGGGATTTACCA 3202
QY 1319 CCCCCGACAAGAGCACCAGAAGGCGCCCTTCCGTGGATGGGTACGAGCTGCACC 1378
Db 3203 CACAGACAAAAACATCAGAAAGACCTCCATTCCTTTGGATGGGTATTGAACCTCATC 3262
QY 1379 CCACAAAGTGGACCGTGCAGCCCATCGAGTGCCTCCGAGAAGGAGAGCTGACCGTGAACG 1438
Db 3263 CTGATAAATGGACAGTACAGCTTAAATGTCGCAGAAAAAGACAGCTGGACTGTCAATG 3322
QY 1439 ACATCAGAAAGCTGGTGGGCAAGCTGNACTGGGCCACCGAGATCTTACCCCGGATCAAGG 1498
Db 3323 ACATACAGAAAGTTAGTGGGAAAATTTGAAATTTGGCAAGTCAAGATTTATGCAAGGATTAAG 3382
QY 1499 TGGCCGAGCTGTGCAAGCTGCTGCGGGCGCCAAAGGCCCTGACCGACATCGTGCCCTGA 1558
Db 3383 TAAAGCAGTTATGTAAACTCCTTAGAGGAACCAAGCCTAACAGAGTAAATACCACATAA 3442
QY 1559 CCAGGAGGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACG 1618
Db 3443 CAGAAGAACGACAGCTAGAACTGCGAGAAAAACAGGGAGATTTCAAAGAAACACAGTACATG 3502
QY 1619 GCGTGTACTAGACCCCGACGAGCTGCTGGCCGAGATCCAGAACGAGGCCACGACC 1678
Db 3503 AAGTATATTATGACCCATCAAAAGACTTAGTAGCAAAATACAGAACGAGGGCAAGGCC 3562
QY 1679 AGTGGACCTACCAAGCTTACCAGGAGCCCTTCAAGAACCTTGAAGCCGGAAGTACGCCA 1738
Db 3563 AATGGACATATCAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGAGAAATGTCAA 3622
QY 1739 AGATGCGACCCCGCCACCAACGACGCTGAAGCAGCTGACCCGAGGCGCTGACAGAAATCG 1798
Db 3623 GGATGAGGGTGGCCACACTAATGATGTAACAGAGTAAACAGAGGCAAGTGCAAAAAGTAT 3682
QY 1799 CCATGAGAGCATCTGGGCGAAGACCCCGAAGTTCCCGCTGCCCTCCCATCCAGAAAG 1858
Db 3683 CCACAGAAAGCATAGTAATATGGGAAAGATTCCTAAATTTAAACTACCCATCAAAAGG 3742
QY 1859 AGACCTGGGAGACTGGTGGACCGACTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1918
Db 3743 AAACATGGGAGCATGGTGGATGGAGTATTGGCAAGCTACTGGATTCTTGGTGGGAGT 3802
QY 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCCATCATCG 1978
Db 3803 TTGTCAATACCCCTCCCTTAGTGAAATTTATGTTACCGTACCAGTTAGAGAAAGAACCCATAGTAG 3862
QY 1979 GCGCCGAGACTTCTAGCTGGAGCGCGCCCGCAACCCCGAGACCAAGATCGGCAAGGCCG 2038
Db 3863 GAGCAGAAATTTCTATGTAGTGGGCGACTAATAGGAGACTAAATTAGGAAAGACAG 3922
QY 2039 GCTACGTGACCGACCGGGCGCGCAGAAAGATCGTAGCCCTGACCGAGACCCACCAACAG 2098
Db 3923 GATATGTACTGACAGAGGAGACAAAAGTTGTCTCCATAGCTGACACAAACAATCAGA 3982
QY 2099 AGACGAGCTGACAGGCAATCCAGCTGGCCCTGCGAGGACACGGCGAGCGAGGTGAACATCG 2158
Db 3983 AGACTGAATTAACAAGCAATTCATCTAGCTTTCGAGGATTCGGGATTTAGAAGTAAACATAG 4042
QY 2159 TGACCGACACGACAGTACCGCCCTGGGATCATCCAGGCCCGCCGACGAGCGAGAGCG 2218
Db 4043 TAAACACTCACAAATATGCATTTAGGAATTCATCAAGCACCAACCCAGATTAAGAGTGAATCAG 4102
QY 2219 AGCTGGTGAACACAGATCATCGACGCTGATCAAGAGGAGAGGTGTACTGTAGCTGGG 2278
Db 4103 AGTTAGTCAGTCAATAATAGACGAGTTAATAAAAAAGGAAGGCTCTACTGTGCACTGGG 4162

QY 1379 CCGACAAGTGGACCGTGCAGCCCATGAGCTGCCGAGAAAGAGAGAGCTGGACCGTGAACG 1438
Db 3263 CTGATAAATGGACAGTACACCTTATATGCTGCCAGAAAGACAGCTGGACTGCTCAATG 3322
QY 1439 ACATCCAGAAGCTGGTGGGCAAGCTGAATCTGGGCCAGCCAGATCTACCCCGGCATCAAGG 1498
Db 3323 ACATACAGAAGTGTAGTGGGAAATTTGAATTTGGGCAAGTCAAGATTTATGCAGGGATTAAG 3382
QY 1499 TGGCCAGCTGTGCAGCTGCTGGCGGCCCAAGCCCTGACCGACATCGTGGCCCTGA 1558
Db 3383 TAAAGCAAGTTATGTAAACTCCTTAGAGGAACCAAGCACCTAACAGAAAGTAAATACCACTAA 3442
QY 1559 CCGAGGAGCGGAGCTGGAGCTGCCGAGAACCGCAGACTCCTCGCGAGCGCCGCTGCACG 1618
Db 3443 CAGAAGAAGCAGAGCTAGAACTGCGAGAAACAGGGAGATTTCTAAAGAACCAAGTACATG 3502
QY 1619 GCGTGTACTACGACCCAGCAAGACCTGTGTGGCCGAGATCCAGAAAGCAGGGCCACGACC 1678
Db 3503 AAGTATATTTATGACCCATCAAAAGACTTTAGTACAGAAATACAGAAAGCAGGGGCAAGGCC 3562
QY 1679 AGTGGACCTACCAAGCTTACCAGAGGCCCTTCAAGACCTGAAGACCGGCAAGTACGCCA 1738
Db 3563 AATGGACATATCAAAATTTATCAGAGGCCATTTAAATCTGAAACACAGAAAGTATGCAA 3622
QY 1739 AGATGCGCACCCGCCACCAACGACGCTGAAGCAGCTGACCGAGGCCCTGCAGAAAGATCG 1798
Db 3623 GGATGAGGGGTGCCACACTAATGATGTAACACAGTTAACAGGGCAGTGCAGAAAGTAT 3682
QY 1799 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCCCAAGTTCCGCCCTGCCCATCCAGAAGG 1858
Db 3683 CCACAGAAAGCATAGTAATATGGGAAAGATTCCTAAATTTAAACTACCCATACAAAAGG 3742
QY 1859 AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
Db 3743 AAACATGGGAAGCATGGTGGATGGAGTATTTGGCAAGCTACCTGGATTCCTGAGTGGGAGT 3802
QY 1919 TCGTGAACACCCCCCTCGTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCGCCCATCATCG 1978
Db 3803 TTGTCATACCCCTCCCTTAGTGAATTTATGGTACCAGTTAGAGAAAGACCCATAGTAG 3862
QY 1979 GCGCCGAGACCTTCTAGCTGGAGCGGCCGCCCAACCGCGAGACCAAGATCGGCAAGCCG 2038
Db 3863 GAGCAGAAACTTTCTATGTAGTGGGAGCTAATAGGGAGACTAAATTTAGGAAAGCAG 3922
QY 2039 GCTACGTGACCGCGGGCGGCGCAGAGATCGTGAGCCTGACCGAGACCAACCAACGAGA 2098
Db 3923 GATATGTTACTGACAGAGGAAGCAAAAAGTTGTCTCATAGCTGACACCAACAAATCAGA 3982
QY 2099 AGACCGAGCTGCAGGGCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCG 2158
Db 3983 AGACTGAATTAACAGCAATTCATCTAGCTTTCAGGATTCGGGATTAGAAGTAAACATAG 4042
QY 2159 TGACCGACAGCCAGTAGTCCCTGGGCATCATCCAGGCCCGCCAGCCAGCAAGAGGAGAGCG 2218
Db 4043 TAAACAGACTCACAAATATGCATTAGGAATCATTCGAACCAACCAAGATAGAGTGAATCAG 4102
QY 2219 AGCTGTTGACCAAGATCATCGACAGCTGATCAAGAGGAGAGAGGTGTACTGAGCTGGG 2278
Db 4103 AGTTAGTCACTCAAAATAAGACAGTAAATAAAAGGAAAGGCTTACCTGGCATGGG 4162
QY 2279 TGCCCGCCCAACAGGGCATCGCGGCAACGAGCAGATTCGACAGCTGGTGAAGCAAGGGCA 2338
Db 4163 TACCAGCACACAAGGAATTTGGAGGAATGAACAAGTAGATAAATAGTCAGTCTCTGGAA 4222
QY 2339 TCCGCAAGGTGCTGTCTCCCTGGACGGCATCGAT 2370
Db 4223 TCAGGAAAGTACTATTTTGAATGGAATAGAT 4254

Db 1266 GATGAAGACTGTACTGAGAGACAGGCTAA-TTTTTAGGAAGATCTGGCCTTCTCTACA 1324
Qy 249 GGGCAAGGCCCGAGTTCCCGAGGAGACAGCCGCGCAACAGCCCGACCGCCGCGA 308
Db 1325 AGGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCAAAGCCCGCCACCAAGAGA 1384
Qy 309 GCTGCAAGTGGGG-----CGACACCCCGCAGGAGCGCGCGCGCGCGCGCGAGG 362
Db 1385 GCTTCAGTCTGGGGTAGACACAACTCCCTCAGAGCAGGAGCCGATAGACNAGG 1444
Qy 363 CA-----CCCTGAACPTCCCGAGATCACCTGTGGCAGCGCCCTGTGTGAGCATCAA 416
Db 1445 AACTGTATCTTTAACTTCCTCAGATCACTCTTTGCAACGACGCCCTCGTCAATAAA 1504
Qy 417 GGTGGCGCCAGATCAAGAGCCCTGCTGGACACCGCGCGCGCGCGAGCACCGTGTGGA 476
Db 1505 GATAGGGGGCAGCTCAAGAGGCTCTCTGGACACCGGAGCAGACACCGTGTGGA 1564
Qy 477 GGAGATGAGCTGCCCGCAAGTGAAGCCCAAGATGATCGGGGCACTCGCGGCTTCAT 536
Db 1565 GGAGATGCTGTGCCAGCGCTGGAAGCCGAAGATGATCGGGGAATCGCGGTTTCAT 1624
Qy 537 CAAGGTGGCCAGTACGACAGATCTGTGATCGAGATCTGGCGCAAGAGCCATCGGCAC 596
Db 1625 CAAGGTGGCCAGTATGACAGATCTCTCATCGAAATCTGGCGCACAGGCTATCGGTAC 1684
Qy 597 CGTCTGATCGGCCCGCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGG 656
Db 1685 CGTCTGATGGGGCGCCACCGCTCAACATCATCGGAGCGAACCCTGTGAGCGAGATCGG 1744
Qy 657 CTGACCCCTGAACCTTCCCATCAGCCCATCGAGCCGCTGCGCCGTGAAGCTGAAGCCCGG 716
Db 1745 TTGCACGCTGAACCTTCCCATCAGACGCTACCGGTGAAGCTGAAGCCCGG 1804
Qy 717 CATGGAGGCCCGAAGTGAAGAGTGGCCCTGACGAGGAGAGATCAAGGCCCTGAC 776
Db 1805 GATGGAGGCCCGAAGTCAAGATGGCCATTCAGAGAGGAGATCAAGGCACTGGT 1864
Qy 777 CGCCATCTCGAGAGATGGAAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCC 836
Db 1865 GGAGATTGACAGATGGAAGAGGAGGAAATCTCCAGATTTGGGCTGAGAACCC 1924
Qy 837 CTACAACACCCCGTGTGGCCATCAAGAAGAGGACAGACCAAGTGGCGCAAGCTGGT 896
Db 1925 GTACAACACCGCGTGTGCGAATCAAGAAGAGGACTCGACGAAATGGCGCAAGCTGGT 1984
Qy 897 GGACTTCGGCGAGTGRACAGCCCGCAGGACTTCTGGAGGTGAGCTGGGCATCCC 956
Db 1985 GGACTTCGGCGAGTGAACAGCGCGCAAGACTTCTGGAGGTTTCACTGGGCGATCCC 2044
Qy 957 CCACCCCGCGGCTGAAGAAGAAGAGCGGTGACCGTGTGGACGTGGGCGACGCTA 1016
Db 2045 GCACCCCGCGGCTGAAGAAGAAGATCGTGACCGTACTGGATGTGGTGATGCTA 2104
Qy 1017 CTTCAAGCGTGGCGAGGACTTCGCAAGTACACCGCTTTCACCATCCCGAGAT 1076
Db 2105 CTTCTCCGTTCCCTGGACGAAGACTTCAGGAAGTACACTGCCCTCACAATCCCTTCGAT 2164
Qy 1077 CAACAACAGACCCCGCATCCGCTACCACTACACGTGTGCTGCCCCAGGCTGGAAGG 1136
Db 2165 CAACAACAGACACCGGGGATTCGATATCACTAACAAGTGTGCCCCAGGCTGGAAGG 2224
Qy 1137 CAGCCCCAGCATCTCCAGACGATCAACCAAGATCTCGAGCGCTTCCCGCGCCGCA 1196
Db 2225 CTCCTCCCATCTCCAGATAGCATGACCAAAATCTTGAGCCCTTTCGCAACAGAA 2284
Qy 1197 CCGCGAGATCTGATCTACAGTACATGGAGCAGCTGTACCTGGGCGAGCAGCTGGAGAT 1256
Db 2285 CCGCGACATCTGATCTATCATGATCATGGATGACTTGTGACCTGGGCTCTGATCTAGAT 2344
Qy 1257 CGGCGACCGCCCAAGATCGAGGAGCTGGCAAGCACCTGCTGCGCTGGGCTTCAC 1316
Db 2345 AGGCGACCGCCCAAGATCGAGGAGCTGGCGCAGCACCTGTTGAGGTGGGCACTGAC 2404

Qy 1317 CACCCCGCAGCAAGAAGCACCAAGAGGAGCCCTTCTCTGTGATGGCTTACGAGCTGCA 1376
Db 2405 CACACCGCAGCAAGAAGCACCAAGAGGAGCCCTTCTCTGTGATGGCTTACGAGCTGCA 2464
Qy 1377 CCGCGCAAGTGGAGCCGTGAGCCCATCGAGCTGCCCGAAGAAGGAGAGCTGGACCTGAA 1436
Db 2465 CCCTGACAAATGGACCTGAGCCCTATCGTGTGCCAGAGAAAGACAGCTGGACTGTCAA 2524
Qy 1437 CGACATCCAGAACTGGTGGCAGCTGAAGTGGCGCAGCAGATCTACCCCGCATCAA 1496
Db 2525 CGACATACAGAACTGGTGGGGAAGTTGAAGTGGCGCAGTCAAGATTTACCCAGGATTA 2584
Qy 1497 GGTGCGCAGCTGTGCAAGCTGTGCGCGCGCAAGCCCTTGACCGACATCTGTCGCCCT 1556
Db 2585 GGTGAGCAGCTGTGCAAACTCTCTCCGCGAACAACGAGCTCACAGAGTGTATCCCT 2644
Qy 1557 GACCGAGGCGCAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCGCTGCA 1616
Db 2645 AACCGAGGCGCAGCTCGAACTGGCAAAACCGAGAGATCTTAAGGAGCCGCTGCA 2704
Qy 1617 CGCGCTCTACTAGACCCAGCAAGGACCTGGTGGCGGAGATCCAGAGCAGGCGCACGA 1676
Db 2705 CGCGGTGTACTATGACCCCTCCAAGGACCTGATCGCGAGATCCAGAGCAGGCGCAAG 2764
Qy 1677 CCAGTGGACCTTACAGATCTTACCAGGAGCCCTTCAAGAACCCTGAAGACCGCAAGTACG 1736
Db 2765 CCAGTGGACCTTACAGATTTTACCAGGAGCCCTTCAAGAACCCTGAAGACCGCAAGTACG 2824
Qy 1737 CAAGATCGCACCCCGCACACCAAGAGCTGAAGACCTGACCGAGGCGCTGCGAGAGAT 1796
Db 2825 CCGATGAGGCGTCCCACTAACGACCTCAAGCAGCTGACCGAGGCGCTGCGAGAGAT 2884
Qy 1797 CGCATGAGAGCATCTGTGCTGGGCAAGACCCCAAGTTCCGCTTCCCATCCAGAA 1856
Db 2885 CACCAACGAAAGCATCTGTGATCTGGGAAAGACTCTTAAGTTCAAGCTGCCCATCCAGAA 2944
Qy 1857 GGAGACCTGGGAGACCTGGTGACCGACTCTGCGAGCCACCTGGATTCCTCCGAGTGGGA 1916
Db 2945 GGAACCTGGGAAACCTGGTGACAGATATTGGCAGGCCACCTGGATTCCTGAGTGGGA 3004
Qy 1917 GTTCGTGAACACCCCGCTGCTGAAGCTGGTGAAGCTGCTGAGAGAGGAGCCCATCAT 1976
Db 3005 GTTCGTGAACACCCCGCTGCTGAGCTGGTGAAGCTGCTGAGAGAGGAGCCCATAGT 3064
Qy 1977 CGCGCGCAGACCTTCTACGTGGAGCGCGCGCAACCGCAGAGACCAAGATCGGCAAGC 2036
Db 3065 GGGCGCGCAACCTTCTACGTGGATGGGCGCGCTAACAGGAGACTAAGCTGGGCAAGC 3124
Qy 2037 CGGTACGTGACCGAGCCGCGCGCGCAAGATCGTGAAGCTGACCGAGACCAACCA 2096
Db 3125 CGGATACGTCACTAACCGGCGCAGACAGAGGTTGTCAACCTCACTGACACCAACCA 3184
Qy 2097 GAAGACCGAGCTGAGGCGCATCCAGCTGGCCCTGCAGACACGCGCAGGCTGAACAT 2156
Db 3185 GAAAGTGAAGTGAAGGCGCATTTTACCTGCTTGGAGGACTCGGCGCTTGGAGGTAACAT 3244
Qy 2157 CGTACCGACAGCAGTACGCGCTGGGCTCATCCAGCGCCAGCGCCGCAAGAGCGAGAG 2216
Db 3245 CGTGACAGCTCTCAGTATGCCCTGGGCTCATTTCAAGCCCGCAGCCAGAGTGAAGT 3304
Qy 2217 CGAGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTG 2276
Db 3305 CGAGCTGGTCAATCAGATCATCGAGCAGCTGATCAAGAAGGAAAGGTTCTATCTGCGCTG 3364
Qy 2277 GGTGCGCGCCAGAGGCGCATCGCGCAAGCAGCAGATCGACAGCTGGTGAAGCAAGG 2336
Db 3365 GGTACCGCGCCAGAGGCGCATTTGGCGCAATGAGCGGTGCGACAAGCTGGTCTCGCTGG 3424
Qy 2337 CATCGCAAGTGTGTTCTCTGGAGCGCATCGA 2369
Db 3425 CATCAGGAAGTGTATTCTCTGGATGCGATCGA 3457

RESULT 2

US-09-872-733-6
; Sequence 6, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-42870S1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBKN containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733-6

Query Match 64.5%; Score 1592.8; DB:10; Length 8366;
Best Local Similarity 80.7%; Pred. No. 2.5e-255;
Matches 1914; Conservative 0; Mismatches 442; Indels 16; Gaps 4;

Qy 14 TGGCGAGGCGCATGAGCGAGCCACAGC---GCCAATCTCTGATGAGCGCAGCAACT 70
Db 1857 TGGCGAGGCGATGAGCGAGCTGACGACTCGGCGNCCATATGATGAGAGGCGCACT 1916
Qy 71 TCAGAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGAGGCGCCACATCGGCC 130
Db 1917 TCCGGAACACAGCGGAAGATCTCAAGTCTTCAATTTGGCAAGAGAGGCGCACACCGCCA 1976
Qy 131 GCAACTGCGCGCGCCCGCGCAAGAGGCTCTGGAAGTGGCGCAAGAGAGGCGCCACACAGA 190
Db 1977 GGAAGTGGCGGCGCGCGCGCAAGAGGCTGTGGAAATGTGGAAAGAGAGGACACAAA 2036
Qy 191 TGAAGGACTGCACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGACTGGCCCTTCCCGCAGG 250
Db 2037 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTATAGGAGAGATGGCCCTTCTTACAG 2095
Qy 251 GCAAGGCGCGGAGTTCCCGAGCGAGCAGACGACCGGCGCAAGCGCCACACGCGCGAGC 310
Db 2096 GGAAGGCGGAGGAATTTCTTTCAGAGCAGACAGAGCCACAGCCCGCCACAGAGAGAGC 2155
Qy 311 TGCAGGTGCGCGG-----CGACAAACCCCGAGGCGCGCGCGCGCGCGCGCGCGCA 364
Db 2156 TTCAGGTCTGGGTAGAGACAAACACTCCCGCTCAGAAAGAGGAGCGCATAGACAAAGAA 2215
Qy 365 -----CCCTGAACCTTCCCGCAGATCACTCTGTGGCAGCGCGCGCGCGCGCGCGCG 418
Db 2216 CTGTATCTTTAACTTCCCTCAGATCACTCTTTGGACAGACCCCTCGTCACATAGGA 2275
Qy 419 TGGCGGCGCAGATCAAGAGGCGCTGTGGACACCGGCGCGAGCACACCGCTGTGGAGG 478
Db 2276 TCGGGGGGCACTCAAGGAGCGCTGCTCGATACAGGAGCAGATGATACATTATAGAG 2335
Qy 479 AGATGAGCCTTCCCGCAAGTGGAGCCAGATGATCGCGCGGATCGCGGCGCTTCATCA 538
Db 2336 AATGAGTTTTCGCGAGAGATGGAAACCAAAATGATAGGGGGATCGGGGGCTTCATCA 2395
Qy 539 AGGTGCGCGAGTACGACAGATCTGTATCGAGATCTCGGCAAGAGGCGCATCGGCACCG 598
Db 2396 AGGTGAGGCGAGTACGACAGATCTCATAGAAATCTGTGGACATTAAGCTATAGGTACAG 2455

Qy 599 TGTGATGCGGCGCCACCGCTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGCT 658
Db 2456 TATTAGTAGGAACTACACCTGTCAACATAATTGGAAGAATCTGTTGACCCAGATCGGCT 2515
Qy 659 GCACCTTGAACCTTCCCATCAGCCCATCGAGACCGCTGCCCGTGAAGCTGAAGCCCGGCA 718
Db 2516 GCACCTTGAACCTTCCCATCAGCCCATTTAGACGCGTCCCGTGAAGTTGAAGCCCGGCA 2575
Qy 719 TGGAGCGCGCCCAAGTGAAGCAGTGGCCCTTACCAGGAGAGAGATCAAGGCGCTGACCG 778
Db 2576 TGGAGCGCGCCCAAGTGAAGCAGTGGCCCTTACCAGGAGAGAGATCAAGGCGCTTAGTGG 2635
Qy 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGCAAGCCCT 838
Db 2636 AATCTGTACAGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCTGAGAACCCCT 2695
Qy 839 ACAACACCCCGTGTTCGCCATCAAGAGAAGAGCAGCACCAAGTGGCCGCAAGCTGGTGG 898
Db 2696 ACAACACTCCAGTCTTCGCAATCAAGAGAAGAGACAGTACCAAGTGGAGAAAGCTGGTGG 2755
Qy 899 ACTTCCGCGAGCTGAACAGCGCCACCGAGGACTTCTGGAGGTGCGAGCTGGGCATCCCC 958
Db 2756 ACTTCAGAGAGCTGAACAGAGAACTCAGGACTTCTGGGAAGTTTCTGAGTGGGCATCCCC 2815
Qy 959 ACCCGCGCGCGCTGAAGAGAAGAGAGGCTGACCGTGTGACGCTGGCGGAGCGCTACT 1018
Db 2816 ATCCGCGTGGGTGAAGAGAAGAGTCACTGACAGTGTGAGTGTGGTGTGCTACT 2875
Qy 1019 TCAGCTGCGCTGGAGGAGACTTCCGGAAGTACACGCGCTTCCACCTTCCAGCATCA 1078
Db 2876 TCTCGTCTCCCTTGGAGCGAGGACTTCAGGAAGTACACTGCTTCCAGTACCTAGCATCA 2935
Qy 1079 ACAAGCAGACCCCGCGCATCCGCTACCAAGTACCAAGCTGCTGCCCGCGAGGCTGGAGGCA 1138
Db 2936 ACAAGCAGACCCCGCGCATCCGCTACCAAGTACCAAGCTGCTGCCCGCGAGGATGGAAGGAT 2995
Qy 1139 GCGGAGCATCTTCCAGAGCAGCATGACCAAGATCTGGAGCGCTTCCGCGCGCGCAAC 1198
Db 2996 CACGAGCATCTTTCAGAGCAGCATGACCAAGATCTTGAGCGCTTCCGCAAGCAAGACC 3055
Qy 1199 CGGAGTCTGATCTTACCATGATGAGGAGCTTACGCTGGCGAGCGACCTTGGAGATCG 1258
Db 3056 CAGACATCTGATCTTATCAGTACATGAGGAGCTTACGCTAGGAGTACCTTGGAGATCG 3115
Qy 1259 GCCAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTCCACA 1318
Db 3116 GCGACACAGACCCAGATCGAGGAGCTGAGACAGCTTGTGTAGGTGGGAGTGAACA 3175
Qy 1319 CCGCGCAAGAGACCAAGAGGAGGAGCGCGCTTCTGTGGATGGCTACGAGCTGCAGC 1378
Db 3176 CACGAGCAAGAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3235
Qy 1379 CCGACAAGTGAACCTGAGCGCATCGAGCTGCGCGAGAGAGAGAGCTGGACCGTGAAGC 1438
Db 3236 CTGACAAGTGAACCTGAGCGCATCGAGCTGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 3295
Qy 1439 ACATCCACAAGCTGTGGGCAAGTGAACCTGGCGCAGCAGATCTACCCCGCATCAAG 1498
Db 3296 ACATACAGAAGCTGTGGGCAAGTGAACCTGGCGCAGCAGATCTACCCAGCATCAAG 3355
Qy 1499 TCGCGCAGCTGTGAAGCTGTGCGCGCGCGCAAGCGCTGACCCGATCTGCTGCCCTTGA 1558
Db 3356 TAGGCACTGTGAAGCTGTGCGCGCGCGCAAGCGCTGACCCAGCATCAAG 3415
Qy 1559 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAGACCGCGAGATCTGCGGAGCGCGCTGACG 1618
Db 3416 CAGAGGAGCAGAGCTAGAACTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3475
Qy 1619 CGGTGTACTACGCGCGAG 1678
Db 3476 GAGTGTACTACGCGCGAG 3535
Qy 1679 AGTGGACCTACAGATCTTACCGAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCA 1738

Db 3536 AATGGACCTACCAATCTACAGAGCCCTTCAAGAACCTGAAGACAGGCAAGTACGCAA 3595
Qy 1739 AGATGCGACCGCCACACACAGAGCTGAAGCAGCTGACCGAGGCGGTGACAGAAATCG 1798
Db 3596 GGATGAGGGGTGCCACACACAGATGTGAAGCAGCTGACAGAGGAGTGCAGAAATCA 3655
Qy 1799 CCATGGAGACATCGTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAGG 1858
Db 3656 CCACAGAGACATCGTGATCTGGGGCAAGACTCCCAAGTTCAAGCTGCCCATACAGAGG 3715
Qy 1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1918
Db 3716 AGACATGGGAGACATGGTGACCGCTACTGGCAAGCCACTGGATCCCTGAGTGGGAGT 3775
Qy 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCACTGGAGAGAGGAGCCCATCATCG 1978
Db 3776 TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTATCACTAGCTGGAGAGGAAACCCATCGT 3835
Qy 1979 GCGCGGAGACCTTCTAGCTGGAGCGCGCGCAACCGCAGACCAAGATCGGAAGSCCG 2038
Db 3836 GAGCAGAGACCTTCTAGCTGGATGGGCGAGCAACAGAGGAGACCAAGCTGGCAAGCAG 3895
Qy 2039 GCTACGTGACCGAGCGCGGCGGCGAGAGATCGTGAAGCTGACCGAGACCAACCAACAGA 2098
Db 3896 GCTACGTGACCAACCGAGGAGCAGAGAAAGTGGTGAAGCTGACTGACACCAACCAACAGA 3955
Qy 2099 AGACCGAGCTGCAGGCCATCCAGCTGCGCCCTGCAGGACACCGCAGCGAGGTGAACATCG 2158
Db 3956 AGACTGAGCTGCAAGCCATTAAGCTAGCTGTGCAAGACACGCGACTGGAAGTGAACATCG 4015
Qy 2159 TGACCGGACGACGCTAGCGCCCTGGGCATCATCCAGGCGCCGAGCGCAAGAGCGAGAGCG 2218
Db 4016 TGACGAGACTCACAGTACGCTGGGCATCATCCNAGCACAACAGCAACCAATCCGAGTCAG 4075
Qy 2219 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGGCTGACTGAGCTGGG 2278
Db 4076 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAGGAGAGGCTGACTGAGCTGGG 4135
Qy 2279 TGCCCGCCCAAGGGGATCGCGGCAACGACGACAGATCGACAGCTGTTGAGCAAGGCA 2338
Db 4136 TACCAGCACACAAAGGAATTTGGAGGAATGAACAAGTAGATAAATAGTCAGTCTGGGA 4195
Qy 2339 TCCGCAAGGTGCTGTTCTCTGACGGCATCGAT 2370
Db 4196 TCCGGAAGGTGCTTCTCTGACGGGATCGAT 4227

RESULT 3

US-09-872-733-1
; Sequence 1, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733-1

Query Match 63.1%; Score 1557.4; DB 10; Length 4338;
Best Local Similarity 80.5%; Pred. No. 1.7e-249;
Matches 1914; Conservative 0; Mismatches 441; Indels 22; Gaps 7;
Qy 14 TGGCCGAGGCGCATGAGCAGGCGCCACCAGC---GCCAACATCTCTGATCAGCGCAGCAACT 70
Db 1085 TGGCCGAGGCGCATGAGCAGGCTGACGAACCTCGCGCACCAATAATGATCAGAGAGGCAACT 1144
Qy 71 TCRAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTCGGCAAGGAGGCGGCACATCGCCCC 130
Db 1145 TCGGAACACCAAGCAAGATCGTCAAGTGTCTTCAATTTGGCAAAAGAGGCGCACACCGCCA 1204
Qy 131 GCAACTCCCGCGCGCCCGCAAGAAAGGGCTGTGGAAAGTGGCGCAAGGAGGCGCCACCAGA 190
Db 1205 GGAATCTCCCGGGCCCCCGGAGAGAGGGCTGTGGAAATGTGGAAGAGGAGGACACCAAA 1264
Qy 191 TGAAGGACTCACCGAGCGCGCAGGCCAACTTCTTCCCGGAGGACCTGGCTTCCCCCAGG 250
Db 1265 TGAAGATTTCTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGGCTTCTCTACAAG 1323
Qy 251 GCAAGGCGCGGAGTTCCCGCAGCGCAGCAGACACCGCCCAACAGCCCGCCAGCGCGAGC 310
Db 1324 GGAAGGCGCAGGGAATTTTCTCAGAGCAGACACAGAGCCCAACAGCCCGCCAGAGAGAGC 1383
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGAGCGCGCGCCAGGCGCA 364
Db 1384 TGCAGGTCTGGGTAGAGACACAACACTCCCGCTCAGAAGCAGGAGCGGATAGACAAGAA 1443
Qy 365 -----CCCTGAACTTCCCGCAGATCACTCTGTGGCAGCGCCCCCTGTGTGAGCATCAAG 418
Db 1444 CTGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCAACAGTAAGA 1503
Qy 419 TGGCGCGCCAGATCAAGGAGCGCTCTGTGGACACCGCGCGCGCAGCAGACCGCTGTGGAGG 478
Db 1504 TCGGGGGGCAACTCAAGGAAGCGCTGCTCGATACAGGAGCAGATGATACAGTATTAGAG 1563
Qy 479 AGATGAGCCTCCCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCA 538
Db 1564 AATGAGTTTGGCAGGAAGATGGAACCAACCAAAATGATAGGGGGATCGGGGCTTCATCA 1623
Qy 539 AGTGCGCCAGTAGCAGCAGATCTCTGATCGAGATCTCGGGCAAGAGCCATCGGCACCG 598
Db 1624 AGTGAGGCGAGTAGCAGCAGATCTCATAGAAATCTGTGGACATATAAGCTATAGTACAG 1683
Qy 599 TGCTGATCGCGCCCGCACC-----CCGCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTG 654
Db 1684 TATTAGTAGGACCTTACCTTACACCTGTCAACATAATTGGAAGAAATCTGTTGACCCAGATC 1743
Qy 655 GGCTGCACTTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGTAAGCTGAAAGCCC 714
Db 1744 GGCTGCACTTGAACCTTCCCATCAGCCCTATTGAGAGCGGTGCCCGTGAAGTGAAGCGG 1803
Qy 715 GGCATGAGCGCCCGCAGGTGAAGCAGTGGCCCTGACCAGGAGAGAGATCAAGGCCCTG 774
Db 1804 GGCATGAGCGCCCGCAGGTGAAGCAGTGGCCCTGACCAGGAGAGAGATCAAGGCCCTTA 1863
Qy 775 ACCGCCATCTCGAGGAGATGGAGAGGAGGCGCAAGATCACCAAGATCGGCGCCCGAGAAC 834
Db 1864 GTCGAAATCTGTACAGAGATGGAGAGGAGGAGATCAGCAAGATCGGCGCTTGAGAAC 1923
Qy 835 CCCTTACAACACCCCGCTGTTCGCCATCAAGAGAGGAGCAGCACAAGTGGCGCAAGCTG 894
Db 1924 CCCTTACAACACTCCAGTCTTCGCAATCAAGAGAGGAGCAGTACCAAGTGGAGAAAGCTG 1983
Qy 895 GTGCACTTCCCGAGCTGAACACAGCGCACCTTCTGGAGGTGCAGCTGGGCATC 954
Db 1984 GTGCACTTCAAGAGCTGAACAGAGAACTCAGGACTTCTGGGAAGTTTCAGCTGGGCATC 2043
Qy 955 CCCACCCCGCGCGCTGAAGAGAGAGCGCTGACCGCTGGAGCTGGCGACGCGC 1014
Db 2044 CCACATCCCGCTGGGTGAAGAGAGAGTCACTGACAGTGTCTGGATGTGGGTGATGCC 2103

Qy	1150	TTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTTCGCGCCGCCCAACCCCGAGATCGTG	1209
Db	367	TTTCAAGCAGCATGACCAAGATCTCTGGAGCCCTTTCGCGAAGCAAAACCCAGACATCGTG	426
Qy	1210	ATCTACCAAGTACATGAGACACCTGTACGTGGGCGAGCGACCTGGAGATCGGCCAGCACCGC	1269
Db	427	ATCTATCAGTACATGAGCAGACCTCTTACCTAGGAAAGTACCTGGAGATCGGGCAGCACAG	486
Qy	1270	GCCAAGATCGAGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACACCCCGGACAAG	1329
Db	487	ACCAAGATCGAGAGCTGAGACAGCATCTGTTGAGTGGGAGCTGACCAACACAGACAAG	546
Qy	1330	AAGCACCAAGAAGAGCCOCCCTTCTGTGGATGGGCTACGAGCTGCAACCCCGACAAGTGG	1389
Db	547	AAGCACCAAGAAGAACCTCCCTTCTGTGGATGGCTACGAACTGCATCTCTGACAAGTGG	606
Qy	1390	ACCGTGCAGCCCATCGAGCTGCCAGAAAGAGAGACTGGACCGTGAACGACATCCAGAAG	1449
Db	607	ACAGTGCAGCCCATCGTGTGCTCGAGAAGACACCTGGAGTGTGAACGACATACAGAAG	666
Qy	1450	CTGGTGGGCAAGCTGAACCTGGGCCACGACATCTACCCCGGCATCAAGGTGGCGACGCTG	1509
Db	667	CTCGTGGGCNAGTTGAACCTGGCAAGCCAGATCTACCCAGGCATCAAGTTAGCAGCTG	726
Qy	1510	TGCAAGCTGCTGGCGGGCGCCCAAGGCCCTCGACGACATCTGTGCCCCCTGACCGGAGAGGCC	1569
Db	727	TGCAAGCTGCTTCGAGGAACCAAGGCACCTGACAGAAGTGATCCCACTTGACAGAGGAAGCA	786
Qy	1570	GAGCTGGAGCTGGCCGAGAACCCGACATCTCTGCCGAGCCGCTGCACGGCTGTACTAC	1629
Db	787	GAGCTAGAACCTGGCAGAGAACCAGAGATCTCTGAAGGAGCCAGTACATGGAGTGTACTAC	846
Qy	1630	GACCCAGCAAGACCTGCTGGCGCAGATCCCAAGACGAGGGCCACGACCACTGGACCTAC	1689
Db	847	GACCCAAAGCAAGCCTGATCGCAGAGATCCCAAGCAGGGCGCAAGGCCAATGACCTAC	906
Qy	1690	CAGATCTACAGAGCCCTTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGGCACC	1749
Db	907	CAAACTTACAGAGCCCTTTCAAGAACCTTGAAGACAGCAAGTACGCAAGGATGAGGGGT	966
Qy	1750	GCCACACCAACGACGTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCGCCATGGAGAGC	1809
Db	967	GCCCAACCAACGATGGAAGCAGCTGACAGAGCGAGTGCAGAAGATCCACAGAGAGC	1026
Qy	1810	ATCGTGATCTGGGGCAAGACCCCAAGTTCCGCTTCCCATCCAGAAGGAGACCTGGAG	1869
Db	1027	ATCGTGATCTGGGGCAAGACTCCCAAGTTCAAGCTGCCATACAGAAGGAGACATGGAG	1086
Qy	1870	ACCTGGTGACCACTACTTGGCAGGCCACCTGGATCCCGCGAGTGGGAGTTCTGTGAACACC	1929
Db	1087	ACATGTGTGACCGAGTACTGGCAAGCCACCTGGATTCCTGTAGTGGGAGTTCTGTGAACACC	1146
Qy	1930	CCCCCTCTGTGAAGCTGTGGTACCAAGCTGGAGAAGGAGCCCATCTATCGGCGCGAGACC	1989
Db	1147	CTTCCCTTGTGAACTGTGGTATCAGCTGGAGAAGGAACCCATCTGTGGGAGCAGAGACC	1206
Qy	1990	TTCTACGTGGACGGCGCCCAACCGAGACCAAGATCGGCAAGCGCGCTACGTGAAC	2049
Db	1207	TTCTACGTGGATGGGCGAGCAACAGGGAGCAACAGCTGGGCAAGGCGGTACGTGACC	1266
Qy	2050	GACCGGGCGGCGAGAAGATCTGTAGCTTGACCGAGACCAACACAGAGACCGGAGCTG	2109
Db	1267	AACCGAGGACGACAGAAGTGTGTGACCTTGACTGACACCAACCAACAGAGACTGAGCTG	1326
Qy	2110	CAGGCCATCCAGCTGGCCCTTGACGACGAGCGGCGAGGTGAACATCTGTACCGCAGC	2169
Db	1327	CAAGCCATCTACCTAGCTGTGCAAGACAGCGGACTGGAAGTGAACATCGTGACAGACTCA	1386
Qy	2170	CAGTACGCCCTGGGCATCATCCAGGCCCAAGCCCGACAAGACGAGAGCGAGCTGGTGAAC	2229
Db	1387	CAGTACGCATGGGCATCATCCAAGCAACAACCAAGCAACCTCCGAGTCAAGGCTGGTGAAC	1446

RESULT 5

US-09-735-487-7

```

US-09-735-487-7
; Sequence 7, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: PATEGGS, Kurt
; APPLICANT: PAWELOS, Kudi
; TITLE OF INVENTION: METHOD OF MANAGING HIV POSITIVE B
; TITLE OF INVENTION: HIV POSITIVE B
; TITLE OF INVENTION: OF HUMAN HIV ST
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/73
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,211
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-735-487-7

```

Query Match 45.8%; Score 1132; DB 10; Length 2601;
Best Local Similarity 68.5%; Pred. No. 3.4e-179;
Matches 1626: Conservative 0; Mismatches 730; Indels 16;

Qy	14	TGGCCGAGGCATGAGCCAGGCCACCA---GCGCCAACATCTGTATGTCAGCCGACGCAACT	70
Db	77	TGGCTGAAGCAATGAGCCAAAGTACAAATTCAGCTACCAATAATGATGTCAGAGAGCAATT	136
Qy	71	TCRAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTTCGCGCAAGGAGGCCACATCGCCC	130
Db	137	TTAGGAACCAAGAAAGATTGTTAAGTGTTTCAATTTGGCAAGAAGGGCACACAGCCA	196
Qy	131	GCAACTTCGCGCGCCCCCGCCGAAGAAGGCGTTCGGAAGTCGCGCAAGGAGGGCCACCAGA	190
Db	197	GAAATTGCAGGGCCCCGTGAAAGAAAGSGCTTTGGAAATGTGGAAAGGAAGGACACCAA	256
Qy	191	TGAAGGACTGACGAGCGGCCAGGCCAACTCTTCCGCGAGGAGACTTGGGCTTCCCCCCAGG	250
Db	257	TGAAGAATTGTACTTGAGAGACAGGCTAA-TTTTTTTAGGGAAGATCTTGGGCTTCCCTACAAG	315
Qy	251	GCAAGGCCCGGAGTTCCTCCAGCGGACGAGAACCGCGCCACACGCCCCACCGCGCGGAGC	310
Db	316	GGAAGGCCAGGGAAATTTCTTCAGAGGACGACCAGAGCCACAGCCCCACCAAGAAGAGC	375
Qy	311	TGCAGGTTCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGCCAGCGGCA	364
Db	376	TTCAGGCTCGGGTAGAGCAACAACATCCCCCTCAGAAGCAGGAGCGGATAGACAAGGA	435
Qy	365	-----CCCTGAACCTCCCGCAGATCACCTGTGGCAGCGCCCTTGGTGAGCATCAAG	418
Db	436	CTGTATCTTTAACTCCCTCAGCTACTCTTTGGCAACGACCCCTCGTCAATAAAGA	495

QY 419 TGGCGGCGCAGATCAAGAGGCGCTTGTGGACACCGGCGCGACGACACCGTGTGGAGG 478
Db 496 TAGGGGGCACTAAAGAAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555
QY 479 AGATGAGCTGCGCGCAGTGGAGGCCAAGATGATCGGCGGCATCGCGGCTTCATCA 538
Db 556 AAATGAGTTTGCAGGAAGATGGAACCAAAATGATAGGGGAATTTGAGGTTTATCA 615
QY 539 AGGTGCGGCATACACCAAGTCTGTGATCGAGATCTGGGGCAAGAGGCCATCGGCACCG 598
Db 616 AAGTAAGACAGTATGATCAGATACATAGATAATCTGTGGACATAAAGCTATAGGTACAG 675
QY 599 TGCTGATCGGCGCCACCGCGTGAACATCATCGGCGGCAACATGTGACCCAGCTGGGCT 658
Db 676 TATTAGTAGGACCTACACTGTCAACATAAATTTGGAAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCTGAACCTTCCCATCATCGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGGCA 718
Db 736 GCACCTTAAATTTCCCATTAGCCCTATTGAGACTGTACCAAGTAAATTTAAAGCCAGGAA 795
QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 796 TGGATGGCCCAAGGTTAAACAATGGCCATTGACAGAGAAATAAAGCATTAGTAG 855
QY 779 CCATCTGCGAGAGATGAGAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 856 AAATTTGTACAGAGATGGAAGAGGAGCAAAATTTCAAAATTTGGGCTGAAATCCAT 915
QY 839 ACAACACCCCGTGTCCGCATCAGAGAGAGGAGCAGACACCAAGTGGCGCAGCTGGTGG 898
Db 916 ACATACTCCAGTATTGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 975
QY 899 ACTTCGGGAGGTGAACAGGCGACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCGCC 958
Db 976 ATTTTCAGAGAACTTAATAGAGAACTCAGACTCTCGGAAGTTCAAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGGCTGAAGAGAGAAAGCGGTGACCGTGTGACGCTGGGCGACGCGCTACT 1018
Db 1036 ATCCCGCAGGCTTAAAGAAAAAATCAGTAACAGTACTGGAATGGGATGATGATAT 1095
QY 1019 TCAGGCTGCCCTGGACGAGGACTTCCCAAGTACAGCCGCTTACCATCCCGACATCA 1078
Db 1096 TTTCAGTTCCCTTAGATCAGAGCTTCAGAGATATCTGCAATTTACCATACCTAGTATAA 1155
QY 1079 ACAACGAGACCCCGGCATCCGCTACCAAGTACAGCTGCTGCCCGAGGCTGGAAGGCA 1138
Db 1156 ACAATGAGACACACAGGGATTAGATATCAGTACAAATGTCTTCCACAGGATGGAAGAT 1215
QY 1139 GCGCCAGCATCTTCAGAGCAGATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACC 1198
Db 1216 CACCAAGCAATATCCAAAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAAATC 1275
QY 1199 CCGAGATCGTATCTACCAAGTACATGAGGAGCTGCGCAAGCAGCTGAGGAGATCG 1258
Db 1276 CAGACATAGTTATCTATCAATACATGGATGATTTGTATGATAGGATCTGACTTAGAAATAG 1335
QY 1259 GCCAGCACCGGCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 1318
Db 1336 GGCAGCATAGAACAAAAATAGAGGAGCTGAGACACATCTGTGTAGGTGGGACTTACCA 1395
QY 1319 CCGGACCAAGAGCACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1378
Db 1396 CACCAAGAAAAACATCAGAGAAACCTCCATCTCTTTGGATGGGTTATGAACCTCCATC 1455
QY 1379 CCGACAAGTGAGCCTGACAGCCATCGAGCTGCGGAGAGGAGGAGGAGGAGGAGGAGG 1438
Db 1456 CTGATAATGAGCAGTACAGCCTATAGTGTGCGCAAAAAAGACAGCTGGAGCTGTCAATG 1515
QY 1439 ACATCCAGAGCTGTGTGGCAAGCTGAATCTGGGCGAGCAGATCTACCCCGGCAATCAAG 1498
Db 1516 ACATACAGAGCTTAGTGGGAAATGANTGGGCAAGTCAGATTTACCCAGGATTAAG 1575
QY 1499 TGCGCCAGCTGTGAAGCTGTGCGCGCGGCAAGGCCCTGACCGACATCTGTGCCCTTGA 1558

Db 1576 TAAGCAATATGTAACTCTTTAGAGGAACCAAAAGCACTAACAGAAGTAATACCACTAA 1635
QY 1559 CCGAGAGGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCGCTGCAGC 1618
Db 1636 CAGAAGAAGCAGAGTAGAACTGGCAGAAAACAGAGAGATTTCTAAAGAAGAACCACTAGTACATG 1695
QY 1619 GCGTGTACTTACCAAGCCCAAGGACCTGTGTGCCAGATCCAGAACGAGTGGGCGGCGGCGG 1678
Db 1696 GASTGTATTATGACCCATCAAGACTTAATAGCAGAAATACAGAAGCAGGGCGCAAGGCC 1755
QY 1679 AGTGACCTTACAGATCTTACCAGGACCTTCAAGAACCTTGAAGACCCGCAAGTACGCCA 1738
Db 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAATACTGAAACACAGGAAAAATGCAA 1815
QY 1739 AGATGGCAGCCCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCG 1798
Db 1816 GAATGAGGGTGGCCACACTATGATGTAACAAATTAACAGAGGCGAGTGAACAAATAA 1875
QY 1799 CCATGGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAGG 1858
Db 1876 CCACAGAAAGCATAGTAATATGGGAAGAGCTCTTAATTTTAACTGCCCATACAAAAG 1935
QY 1859 AGACCTGGGAGACCTTGTGGACCGCTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1918
Db 1936 AAACATGGGAAACATGTTGGACAGAGTATTGGCAAGCCACCTGGATCTCTGAGTGGGAGT 1995
QY 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGGAGGAGGAGGAGT 1978
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGTTACAGTTAGAGAAAGAACCCATAGTAG 2055
QY 1979 GCGCGAGACCTTCTACGTGGAGCGCGCCCAAGCGGAGACCAAGATCGGCAAGGCGG 2038
Db 2056 GAGCAGAAACCTTCTATGTAGATGGGCGAGCTTACAGGAGACTAATTAGGAAAAGCAG 2115
QY 2039 GCTAGCTGACCGACCGCGGCGCGCAGAGATCGTGAAGCTTACCGGAGCAGCAGCAGCAG 2098
Db 2116 GATATGTTACTAATAGAGGAAGACAAAAAATTTGTACCCCTAAGTACACAAACAAATCAGA 2175
QY 2099 AGACCGAGCTGAGCGCCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGAGGAGGAGT 2158
Db 2176 AGACTGAGTTTACAGCAATTTATCTAGCTTTGAGGATTTGAGAAAGTAAACATAG 2235
QY 2159 TGACGAGACCGCAGTACGCTTGGCATCATCCAGGCCAGCGCCAGAGAGAGAGAGAGG 2218
Db 2236 TAACAGACTCAAAATATGATTAGGAATCTTCAAGCAGCAACAGATCAAGTGAATCAG 2295
QY 2219 AGCTGGTGAACAGATCATCTGAGCAGCTGATCAAGAAGGAGAGGAGTGTACCTGAGCTGG 2278
Db 2296 AGTTAGTCAATCAAAATATAGAGCAGTTAATAAAAAAGGAAAGGTCTATCTGGCATGG 2355
QY 2279 TGCCCGCCCGCAGGCGCATCGCGCAGCAGCAGCAGATCCAGAGCTGGTGGAGGAGGCA 2338
Db 2356 TACCAGCAGCAAGGAAATTTGAGGAAATTAAGCAAGTAGATAAATAGTCAAGTGTCTGGAA 2415
QY 2339 TCCGCAAGTGTCTTCTGGAGCGCATCGAT 2370
Db 2416 TCAGGAAGTACTATTTTTAGATGGAATAGAT 2447

RESULT 6

US-09-735-487-9

; Sequence 9, Application US/09735487

; Patent No. US20020042679A1

; GENERAL INFORMATION:

; APPLICANT: de BETHUNE, Marie-Pierre

; APPLICANT: HERTOGS, Kurt

; APPLICANT: PAUWELS, Rudi

; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE

; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY

; FILE REFERENCE: 1377-125P

; CURRENT APPLICATION NUMBER: US/09/735,487

; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)..(489)
; OTHER INFORMATION: gag P6 (52 AA)
US-09-735-487-9

Query Match 45.8%; Score 1132; DB 10; Length 2601;

Best Local Similarity 68.5%; Pred. No. 3.4e-179;

Matches 1626; Conservative 0; Mismatches 170; Indels 16; Gaps 4;

```
QY 14 TGGCCGAGGCGCATGAGCGAGGCCACCA--GCGCCAACATCTCTGATGCGAGCGAGCAACT 70
      ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 77 TGGCTGAGCAATGAGCCAGTAACAANTTCAGTACCATAATGATGCAGAGGCAATT 136
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 71 TCAAGGCCCCCAAGCGCATCAAGTGTCTCAACTGCGGCAAGGAGGCCACATCGCC 130
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 137 TTAGGAACCAAGAAAGATTGTTAAGTGTTCATTTGTCGAAGAAAGGCGCACAGCCA 196
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTCCGCAAGGAGGCGCACCA 190
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 197 GAATTTGACGGGCCCTTAGGAAGAGGGCTGTTGGAATATGGAAGGAAGGACACAAA 256
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 191 TGAAGGACTCACCGAGCGCGGCGCAACTCTTCCGCGAGGACCTTGGCCCTTCCCGCAGG 250
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 257 TGAAGATTCTAGAGACAGGCTAA-TTTTTAGGGAAGATCTGGCCCTTCTTACAAG 315
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 251 GCAAGGCCCGGAGTTGCCCGAGCGAGAGAACCGCGCCACAGCCGCCACAGCGCGAGC 310
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 316 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACACAGAGCCACAGCCGCCACCAAGAGAGC 375
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 311 TGCAGGTGCCGG- ----CGACAAACCCCGCAGCGAGGCGCGCGCGAGCGCGAGGCA 364
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 376 TTCAGGTCTGGGTAGAGACAACAATCCCGCTCAGAGCAGGAGCGGATAGACAGGAA 435
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 365 -----CCCTGAATTCGCCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGG 418
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 436 CTGTATCTTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCGTCAACAATAAGA 495
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 419 TGGGCGCCAGATCAGAGGAGCCCTGCTGGACACCGCGCGCGAGCACCGCTGCTGAGG 478
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 479 AGATGAGCCTCGCCGCACTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 556 AATGAGTTTGGCAGGAGATGGAACCAAAATGATAGGGGAAATTTGAGGTTTTATCA 615
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 539 AGGTGCGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 616 AAGTAACACAGTATGATCAGATCTCATAGAAATCTGTGGACATAAAGCTTATAGGTACAG 675
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 599 TGTGATGCGGCCACCCCGTGAACATCATCGGCGGCAACATGCTGACCCAGCTGGGCT 658
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 676 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTT 735
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 659 GCACCTGGAATTTCCCATCAGCCCATCAGACCGTGCCTGCGCTGAAGCTGAAGCCCGCA 718
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 736 GCACTTTAAATTTTCCCATTTAGCCCTATTGAGACTGTACCGTAAATTAAGCCAGGAA 795
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 796 TGGATGCCCAAAAGTTAAACAATGGCCATTGACAGAGAAATAAAGCATTAGTAG 855
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 779 CCATCTCGAGAGATGGAGAGGAGGCAAGATCACCAAGATCGCGCCCGCGAGAACCCCT 838
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


Db 556 AATAGTTGTCAGGAGATGGAACCAAAATGATAGGGGAATTGGAGGTTTTATCA 615
QY 539 AGTGGCCAGTACGACCATCTGATCGAGATCTCGCCGCAAGAGGCCATCGGCACCG 598
Db 616 AAGTAAGACAGTATGATCATGACTACTATAGAAATCTGTGGACATAAGCTATAGGTACAG 675
QY 599 TGTGATCGGCCACCCCGTGAACATCATCGCGCGCAACATCTGTGACCCAGCTGGCT 658
Db 676 TATTAGTAGACCTACACCTGTCAACATAATTGGAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCTGTAACCTTCCCATCAGCCCATCGAGACCTGTGCCGTGAGCTGAGCCCGGCA 718
Db 736 GCACCTTAAATTTCCCATTTAGCCCTATTGAGACTGTACCACTGTAATAAATAAGCCAGAA 795
QY 719 TGCACGCCCAAGGTGAACGAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 796 TGGATGCCCAAGAGTTAAACAATGGCCATTGACAGAGAAATAAAGCATATTAGTAG 855
QY 779 CCATCTGCGAGAGATGGAGAGGAGGCAAGATCAACCAAGATCGGCCCGGAGAACCCCT 838
Db 856 AATTTGTACAGAGATGGAAGAGGAGCAAAATTTCAAAATTTGGCCTGAAATCCAT 915
QY 839 ACAACACCCCGTGTGCGCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTG 898
Db 916 ACAATACCTCCAGTATTGGCCATAAGAGAAAGACAGTACTAAATGGAGAAATTAGTAG 975
QY 899 ACTTCGCGAGCTGAACAGCGCACCCAGCACTTCTGGGAGGTGCAGCTGGCATCCCC 958
Db 976 ATTTAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAATTTCAATAGGAATACAC 1035
QY 959 ACCCGCGCCCTGAAGAGAGAGAGCGTGCCTGCTGGAGCTGGCGGAGCGCTACT 1018
Db 1036 ATCCCGAGCTTAAAGAGAAATACTAGTACAGTACTGGATGGGTGATGCATATT 1095
QY 1019 TCAGTGTCCCTGGAGAGGATTCGGAAGTACACCGCTTCACCATCCCGAGCATCA 1078
Db 1096 TTTCACTTCCCTTAGATGAAGACTTCAGGAAGTATAGTGCATTTACCATACCTAGTATA 1155
QY 1079 ACAACAGACCCCGCATCCGCTACAGTACACAGTGCCTGCCAGGCGTGGAGGGCA 1138
Db 1156 AATAGAGACACAGGATTAGATATCAGTACAAATGTGCTTCCACAGGATGGAAGGAT 1215
QY 1139 GCGCCAGCATCTTCCAGAGCAGATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACC 1198
Db 1216 CACCAGCAATTTCCAAAGTAGCATGACAAATACTTTAGAGCCTTTAGAAACAAATC 1275
QY 1199 CCGAGATCGTGTATCCAGTATAGTACGACCTGTACGTGGGAGGAGCGCTGGAGTCG 1258
Db 1276 CAGACATAGTTATCTATCAATAGATGGATGTTGTATGTAGGATCTGACTTAGAATAG 1335
QY 1259 GCAGACACCGCCAGATCGAGAGCTGCGCAAGCACTGCTGCGCTGGGCTTCACCA 1318
Db 1336 GGAGCATAGAACAAATAGAGAGCTGAGACAACTCTGTGAGTGGGACTTACCA 1395
QY 1319 CCGCCGCAAGAGACACAGAGAGAGCCCTCTCTGTTGGATGGCTACGAGTGCACC 1378
Db 1396 CACCAGACAAACATCAGAAAGAACCTCCATTCTCTTGGATGGGTATGAATCCATC 1455
QY 1379 CCGAAGTGGAGCTGCGACCCCATGAGTGCCTCGGAGAGAGAGCTGGACCTGAGC 1438
Db 1456 CTGATTAATGGAGTACAGCTTATAGTGTGCGCAAGAGAGAGCTGGACTGTCAATG 1515
QY 1439 ACATCCAGAGAGCTGGTGGGAGCTGAACCTGGCCAGCCAGATCTACCCCGCATCAAG 1498
Db 1516 ACATACAGAGAGTGTAGTGGGAATTTGAATTTGGGCAAGTCAAGTATCCAGGGATTAAG 1575
QY 1499 TGGCCAGCTGTGCAAGCTGCTGCGGGCGCCCAAGGCCCTGACCGACATCGTGGCCCTGA 1558
Db 1576 TAGGCAATTTATGTAACCTCTTAGAGAGACCAAGACCTAACAGAGTATATACCATAA 1635
QY 1559 CCGAGAGCGCGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGCGGAGCCCGTGCACG 1618

Db 1636 CAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTTAAAAGAACCACTACATG 1695
QY 1619 GCCTGTACTACGACCCCGCAGCAAGCACTGCTGGCCGAGATCCAGAAGCAGGCGCACACC 1678
Db 1696 GAGTGTATTATGACCCATCAAGAGACTTAATAGCAAAATACAGAGCAGGGGCAAGGCC 1755
QY 1679 AGTGGACCTACCAAGATCTACAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1738
Db 1756 AATGGACATATCAAAATTTATCAAGAGCCATTTAAATCTGAAAACAGGAAATATGCAA 1815
QY 1739 AGATGGCCACCGCCACACCAAGACGCTGAAGCAGCTGACCGAGGCGCTGAGAGATCG 1798
Db 1816 GAATGAGGGTGGCCACACTAATGATGTAACAAATTAACAGAGGCACTGCAAAAAATAA 1875
QY 1799 CCATGGAGACATCTGATCTGGGGCAAGACCCCAAGTTCCCGCTGCCATCCAGAGG 1858
Db 1876 CCACAGAAACATAGTAATATGGGAAAGACTCTCTAAATTTAAACTGCCATACAAAGG 1935
QY 1859 AGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCACTTGGATCCCGAGTGGAGT 1918
Db 1936 AAACATGGGAACATGCTGGACAGAGTATTGGCAAGCCACTGGATTCCTGAGTGGAGT 1995
QY 1919 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCG 1978
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGTTACCAAGTTAGAGAAAGAACCCATAGTAG 2055
QY 1979 GCGCCGAGACCTTCTACGTGGCGCGCCGCAACCGCAGAGCAAGATCGGCAAGCCG 2038
Db 2056 GAGCAGAAACCTTCTATGTAGTGGGCGAGCTAACAGGGAGACTTAATAGGAAAGACAG 2115
QY 2039 GCTACGTGACCGAGCGGGCGGCGAGAGATTCGTGAGCCCTGACCGAGACCAACCAACAGA 2098
Db 2116 GATATGTTACTATAGAGGAGACAAAAAGTTGTCACTTAACCTGACACAAATCAGA 2175
QY 2099 AGACCGAGCTGAGGCGCATCCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCG 2158
Db 2176 AGACTGAGTTACAGCAATTTATCTAGCTTTGCAAGATTCGGGATTTAGAACTAAACATAG 2235
QY 2159 TGACCGACACCACTGCTGCGCATCTCCAGGCGCCGAGCCGCAAGAGCAGAGCG 2218
Db 2236 TAACAGACTCACATATGTCATTTAGTAATCATTTCAAGCACACCAAGATCAAGTGAATCAG 2295
QY 2219 AGCTGTGAAACCAAGATCTGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2278
Db 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAAGGAAAGGCTCTATCTGCGCATGG 2355
QY 2279 TGCCCGCCACAGGGCATCGCGGCAACGAGCAGATCGCAAGCTGCTGAGCAGGCA 2338
Db 2356 TACCAGCACAAAGGAATTTGGAGAAATGAACAAGTAGATAAATTAGTCAGTCTGGAA 2415
QY 2339 TCCGCAAGGTGCTGCTTCTTCCGACGCGCATCGAT 2370
Db 2416 TCAGGAAGTACTATTTTAGATGGAATAGAT 2447

RESULT 9

US-09-183-1
; Sequence 1, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 1
; LENGTH: 4307

Db	3127	GATATGTTACTAATAGAGGAACAAAAAGTTGTCAACCTTAAC	TGCACACAAACAATCAGA	3186
QY	2099	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGGGCAGGAGGTGAACATCG	2158	
Db	3187	AGACTGAGTTACAAGCAATTTATCTAGCTTTGCGGATTCGGATTTAGAAGTAAACATAG	3246	
QY	2159	TGACCCACAGCCAGTACGCCCTGGGCATCATCCAGGCCGCCGCACNAGGCGAGGCG	2218	
Db	3247	TAAAGACTCACATATGCAATTAGGAATCATTTCAAGCACACCCAGATCAAGTGAATTCAG	3306	
QY	2219	AGCTGTGTAAACCAGATCATCGAGCAGCTGATCAAGAAGGGAAGGTGTACCTGAGCTGGG	2278	
Db	3307	AGTTAGTCAATCAAAATAATAGAGCAGTTTAATAAAAAAGGAAAAGTCTATCTGCCATGGG	3366	
QY	2279	TGCCCCCCACAGGSGCATCGCGGCCACAGCAGCAGATCGACAAGCTGGTGAGCAGGGCA	2338	
Db	3367	TACACGCACACAAAGGAATTTGGAGGAATTGAACAAGTAGATAAATTTAGTCAAGTCTGCTGGAA	3426	
QY	2339	TCGCGAAGGTGCTGTTTCCTGGAGCGGCATCGAT	2370	
Db	3427	TCAGGAAAGTACTATTATTTTATAGATGGAATAGAT	3458	

RESULT 10

RECEIVED
US-09-737-190A-1

Sequence 1, Application US/09737190A

; Patent No. US20020102545A1

; GENERAL INFORMATION:

APPLICANT: Shibuya, Tetsuo

; TITLE OF INVENTION: A Method for Changing a Target Array, a Method for Analyzing

TITLE OF INVENTION: a Structure, and an Apparatus, a Storage Medium and a

; TITLE OF INVENTION: Transmission Medium Therefor

; FILE REFERENCE: JP919990270US1 (14043)

; CURRENT APPLICATION NUMBER: US/09/737,190A

; CURRENT FILING DATE: 2000-12-14

; NUMBER OF SEQ ID NOS: 2

```

; SOFTWARE: PatentIn ver. 2.1

```

; SEQ ID NO 1

```

; LENGTH: 9719
; TYPE: DNA

```

; TYPE: DNA
;
; : Human ; Tm=60.0°C ; GC=49.8% ; Size=107 bp

US-00-737-190A-1
; ORGANISM: Human immunodeficiency virus type 1

US-09-737-I90A-1

Query Match	45.8%	Score 1132;	DB 10;	Length 9719;
Best Local Similarity	68.5%	Pred. No. 3.6e-179;		
Matches 1626;	Conservative	0;	Mismatches 730;	
			Indels 16;	Gaps 4;

Qy	14	TGCGGAGGCCATGAGCCAGGCCACCA---GGGCCAACATCCTGTGATGAGCGCAGCAACT	70
Db	1877	TGGCTGAAGCAATGAGCCAAAGTAACAAATTCAGCTACCATATGATGACAGAGGCAATT	1936
Qy	71	TCAAGGGCCCCAAGCCGATCATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACATCGCCC	130
Db	1937	TTAGGAACCAAGAAAGATGTTAAAGTGTTCATTTGCGCAAGAAGGGCACACAGCCA	1996
Qy	131	GCAACTGCCGCGCCCCGCCAGAAAGGGCTGCTGAACTGGCGCAAGGAGGCCACCAGA	190
Db	1997	GAAATTCGAGGGCCCCTAGGAAAAAGGGCTGTGAAATGTGGAAGAAGGAGCACCAAA	2056
Qy	191	TGAAGGACTCACCGAGCGCAGGCCAACTCTTCGCGGAGGACTGSCCTTCCCCCAGG	250
Db	2057	TGAAGATTGTACTGAGAGACAGGCTAA-TTTTTTTAGGAAGATCTGCGCTTCTTACAAG	2115
Qy	251	GCAAGGCCCGCAGTTTCCCGACGAGCAACCGCGCCAAAGCCCAACACCGCGCGAG	310
Db	2116	GGAGGCCAGGGATTTTCTTCAGAGCAGCAGCAGCCACAGCCCCACACAGAGAGAGC	2175
Qy	311	TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGCGCGCGCGAGGCCAGGGCA	364
Db	2176	TTCAGGTCTGGGTAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCCGATAGACAAGAA	2235
Qy	365	-----CCCTGAACTTCCCGCCAGATCACCTGTGCGAGCGGCCCTGGTGGACATCAAGG	418

```
QY 1499 TGGCCAGCTGTGCAAGCTGCTGGCGCGCGCCAGGCGCCCTGACCGACATCGTGCCCTGA 1558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3376 TAGGCAATTATGTAAGTCTCTTAGAGAGACCAAGACCTAACAGAAAGTAATACCACTAA 3435
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1559 CCGAGAGCGCGAGCTGGAGCTGGCCAGAGACCCGCGAGATCTTGGCGAGCCGCTGACG 1618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3436 CAGAAGAGCAGAGCTAGAAGCTGGCAGAAAACAGAGAGATTCTAAAAGAACCAATACG 3495
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1619 GCGTGTACTACGACCCAGCAAGGACCTGTGGCGCGAGATCCAGAGCAGGCGCACGACC 1678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3496 GAGTGTATTATGACCCATCAAGACTTAATAGCAGAAATACAGAGCAGGGCGAAGGCC 3555
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1679 AGTGGACCTACAGATCTACGAGGACCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3556 AATGGACATATCAAAATTATCAGAGCCATTAAAAATCTGAAAACAGGAAATATGCAA 3615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1739 AGATGCCACCGCCACACACGACGCTGAAGCAGCTGACCGAGGCGCGTGCAGAAATCG 1798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3616 GAATGAGGGGTGCCACACTAATGATGTAACAATTTAAGAGAGGCGAGTGCAAAAATAA 3675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1799 CCATGGAGAGCATCGTATCTGGGGCAAGACCCCAAGTTCCGCTCGCCCATCCAGAGG 1858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3676 CCAGAGAACGATAGTAATATGGGAAAGACTCTTAATTTAACTGCCCATCAAAAGG 3735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1859 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACTGGATCCCGAGTGGGAGT 1918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3736 AAACATGGGAAACATGGTGACAGAGTATTGGCAAGCCACTGGATTCTCTGAGTGGAGT 3795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1919 TCGTGAACACCCCGCTCGTGAAGCTGTGTACAGCTGGGAGAGGAGCCCATCATCG 1978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3796 TTGTTAATACCCCTCCCTTAGTGAATTTATGGTACCAGTTAGAGAAAGAACCCATAGTAG 3855
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1979 GCGCCGAGACCTTCTAGCTGGAGCGCGCGCCCAACCGCGAGACCAAGATCGCAAGCGCG 2038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3856 GAGCAGAACTTCTATGTATAGTGGCGACGCTACAGGGAGACTAATTTAGGAAAGCAG 3915
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2039 GCTACGTGACCGACCGGGCGGCGAGAGATCGTGAAGCTTGACCGAGACCAACCCAGA 2098
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3916 GATATGTTACTAATAGAGGAAGACAAAAGTTGTACCCCTTAAGTGACACAAACAAATCAGA 3975
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2099 AGACCGAGCTGCAGGCGATCCAGCTGCGCTGCAGGACACGCGGAGGTGAACATCG 2158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3976 AGACTGAGTACAGGCAATTTATCTAGCTTTGAGGATTGCGGATTAGAAATTAACATAG 4035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2159 TGACCGACACGCCAGTACGCGCTGGGATCATCCAGGCGCCAGCCGACAGAGCGAGAGCG 2218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4036 TAACAGACTCACAATATGCATTAGGAATCATTCAGCACAACCCAGATCAAGTGAATCAG 4095
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2219 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4096 AGTTAGTCAATCAAAATAAGAGCAGTTAATAAAGGAAAAGGTTCTATCTGGCATGGG 4155
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2279 TGCCCGCCCAAGGGCATCGGCGGCAACAGCAGAGCTGACAGAGCTGGTGAGCAAGGGCA 2338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4156 TACCAGCACACAAAGGAATTTGGAGGAAATGAACAGTAGAATTAATAGTCAGTCTGGAA 4215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2339 TCAGCAAGTGTCTTCTCGGCGGCGATCAT 2370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4216 TCAGGAAGTACTATTTTAGATGGATAGAT 4247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

```
US-09-943-286-3
; Sequence 3, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; OTHER INFORMATION: plasmid.
US-09-943-286-3
```

```
Query Match 43.7%; Score 1079.6; DB 10; Length 8933;
Best Local Similarity 55.7%; Pred. No. 1.7e-170;
Matches 1341; Conservative 281; Mismatches 734; Indels 52; Gaps 5;
```

```
QY 14 TGCGCGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTCTGATGCAGGCGAGCAACT 70
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1200 UGCUGAAGCAUUGAGCCAAAGAACAAUAUACGUACCAUAAUUGAUGCAGAGAGGCAAU 1259
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTCAACTCGGCAAGGAGGCGGCACATTCGCC 130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1260 UUAGGAACCAAGAAAGAUUGUUUAAGUUUAUUGGCAAGAAAGGCGCACACAGCCA 1319
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 131 GCAACTCGCGCGCCCGCCCGCAAGAAAGGCTCTGGAAGTGCGGCAAGAGGGCCACCAGA 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 GAAAUUGCAGGGCCCVAGGAAAAGGCGUGUUGGAAAUGUGGAAAGGAAAGGACACAAA 1379
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 TGAAGGACTCCAGCGCGCCAGCCCAACTTCTTCCCGGAGGACCTGGCTTCCCCCAGG 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 UGAAAGAUUGUACUGAGAGACAGGCUAA-UUUUUUAGGGAAGAUUCUGGCCUUCUACAAG 1438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 251 GCAAGGCGCCCGCGAGT-----TCCCCAGCG 274
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1439 GGAAGCGCCAGGAUUUUUUCAGAGCAGACCAGCCACAGCCCCACCAUUUCUUCAG 1498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 275 AGCAGAACCCGCGCAACAGACCCACCAGCCGCGAGCTGCAGTGCAGTGCAGG-----CGACA 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1499 AGCAGACCCAGGCAACACGCCCCACCAGAGAGCUUACAGGUCUGGGUGAGACAACA 1558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 329 ACCCGCCAGCGCGCGCGCGCGCGCGCGCGCA-----CCCTGAACCTTCCCCCAGA 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1559 ACUCCCCUCAGAGCAGGAGCGCGGAUAGACAAGAACUGUAUCCUUUAUCCUCCUGA 1618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 TCACCCCTGTGCGAGCGCCCTGTGTAGCATCAAGTGGCGCGCCAGATCAAGAGAGGCC 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1619 UCACUUUUGGCAACGACCCCGUCACAAUAAAGAUAGGGGGCAACUAAAGAAAGCUC 1678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 TGTGTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1679 UAUUAGAUACAGGAGCAGAUACAGAUUUAAGAAAUAGAUUUGCCAGGAAGAU 1738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 503 AGCCCAAGATGATCGCGCGCATCGCGGCTTCATCAAGTGGCGCGCGCGCGCGCGCGCG 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1739 AACCAAAAUAGUAGGGGAAUUGGAGUUUAUCAAGAAUAGACAGAUUAGUACAGAUAC 1798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 563 TGATCGAGATCTCGGCAAGAGGCCCATCGGCACCGTGTGATCGGCCCGCCACCGCCCTGA 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 UCAUAGAAUUCUGUGGACAUAAAGCUUAGUACAGAUUUAUAGUAGGACCUACACCUUCA 1858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 ACATCATCGCGCGCAACATCTGACCCAGCTGGGCTGCAACCTGAACTTCCCCATCAGCC 682
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 ACAUAAUUGGAAGAAUUCUGUACUCAGAUUGGUUGCAUUAUUUUUUUUUUUUUUUAGCC 1918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 CCATCAGACCGTGGCGTGAAGCTGAAGCCCGCGCATGGAGCGGCCCGCCCAAGTGAGAGT 742
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CUUUGAGACUGUACCAAGUAAAAUUAAGCCAGGAUUGGAGGCCCAAAAAUUUAAAAAU 1978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 743 GGCCCTTGACCGAGGAGAAAGTCAAGCCCTGACCGCCATCTGCGAGGAGATGGAGAGG 802
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 GGCAUUGACAGAGAAAAAUAAGCAUUAUAGUAGAAAUUUGUACAGAAAUUGAAAGG 2038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 803 AGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACAACACCCCGCTGTTCGCCATCA 862
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 1320 GAAUUGCAGGGCCCCUAGGAAAGGCGUGUUGGAAAUUGGAAAGGAAGGACACAAA 1379
Qy 191 TGAGGAGTACCGAGCGCCAGGCCAATCTTCGCGAGGAGCTGGCCCTCCCCAGG 250
Db 1380 UGAAGAUUGUGAGAGAGAGGCUAA-UUUUUUAGGGAAGUUGGCGCUUUCUACAG 1438
Qy 251 GCAAGGCCCCGAGT-----TCCCCAGG 274
Db 1439 GGAAGGCCAGGAUUUUUUCAGAGCAGACAGCCACAGCCCCACCAUUUCUACG 1498
Qy 275 AGCAAAACCGGCGCAACAGCCCGAGCGGAGCTGCGAGTGCAGGCGCGG-----CGACA 328
Db 1499 AGCAGCAGAGGCAACAGCGCCACCAAGAGAGAGCUUAGGUCUGGGUAGAGACAA 1558
Qy 329 ACCCCGAGGAGGCGCGCGAGCGCCAGGCA-----CCCTCAACTTCCCCAGA 382
Db 1559 ACUCCCCCUAGAAGCAGGAGCGGAUAGCAAGGAACUGUAUCCUUUAUCCUCCAGA 1618
Qy 383 TCACCTGTGGCAGCGCCCTGTGTAGCATCAAGTGGGCGGCAGATCAAGGAGGCC 442
Db 1619 UCACUUCUGCAACGACCCUCGUCACAAUAAAGUAGGGGCAACUAAAGGAAGCUC 1678
Qy 443 TGCTGGACAGCGCGCGAGCAGACACCTGTGTGGAGGAGATGAGCTCCCGCGCAAGTGA 502
Db 1679 UAUAGAUACAGGAGAGAGUACAGAUUAGAGAAUAGUUUUGCCAGGAAGUGGA 1738
Qy 503 AGCCCAAGATGATCGCGGATCGCGGCTTCATCAAGTGCAGCTAGACAGATCC 562
Db 1739 AACCAGAAUAGUAGGGGAUUGGAGUUUUUUAAGUAAGACAGAUUGAUCAGAUAC 1798
Qy 563 TGATCGAGATGTGCGGCAAGAGCCATCGCCACCGTGTGATCGGCGCCACCCCGTGA 622
Db 1799 UCAUAGAAUUCUGGACAUAAAGCUUAGGUAUAGUAGGACCUACACCUUGCA 1858
Qy 623 ACATCATCGGCGCAACATGCTGACCGAGCTGGGCTGCACCTCAACTTCCCATGAGCC 682
Db 1859 ACAUAAUUGGAAAGAAUUCUGUACUAGUUGGUCUUAUUAUUUCCCAUAGCC 1918
Qy 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCCGCATGAGCGGCCCAAGGTGAAGAGT 742
Db 1919 CUUUGAGAGUACAGUAAAUUAAAGCAGGAUUGGAGGCGCCAAAGUUAAACAU 1978
Qy 743 GGCCCTGACCGAGGAGAAATCAAGCCCTGACCGCATCTGCGAGGAGATGAGGAAG 802
Db 1979 GGCCAUUGACAGAGAAAUAAAGCAUUAUAGAAAUUUUUGACAGAAUUGGAAAGG 2038
Qy 803 AGGCAAGATCACCAAGTCCGCGCCCGGAGACCCCTACACACCCCGTTCGCCATCA 862
Db 2039 AAGGAAAUUUCAAAAUUGGCGCUGAGAAUCCAUCAUACUCCAGAUUUUGCCAUAA 2098
Qy 863 AGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGTGGACTTCGCGAGCTGAACAGGCA 922
Db 2099 AGAAAGAGACAGUACUUAUGGAGAAAUUAGUAGAUUUCAGAGACUUAUAAAGAA 2158
Qy 923 CCCAGGACTTCTGGAGGTGCGAGTGGGATCCCCCAACCGCGCGCTGGAAGAGAA 982
Db 2159 CUCAAGACUUCUGGGAUUAUUAUAGAAUACACAUUCCGCGAGGUUAAAAAGAAA 2218
Qy 983 AGAGCTGACCGTGTGAGCTGGGCGAGCGCTACTTCAGGTGCGCTGACGAGACT 1042
Db 2219 AAUCAGAAACAGUAGUUGGUGUAGUAGUUAUUUUUAGUUCUUUAGUAGAGACU 2278
Qy 1043 TCCCAAGTACACCGCTTCACTACCTCCCGAGCATCAACAGAGACCCCGCGCATCGCT 1102
Db 2279 UCAGGAAGUACUGCAUUUACCAUACCUAGUAUAAACUAGACACACCGGAUAGAU 2338
Qy 1103 ACCAGTACAGCTGTGCGCCAGGCTGGAAGGAGCGCCCGCAGCATCTTCAGAGCAGCA 1162
Db 2339 AUCAGUACAAUUGUCCACAGGAGUAGGAAGGAUACACAGCAAAUUAUCCAAAGUAGCA 2398
Qy 1163 TGACCAAGATCTTGAGGCGCTTCCGCGCGCCCAACCGCGAGATCGTGTATCTACAGTACA 1222
Db 2399 UGACAAAUAUUCUAGAGCCUUUUUUAAAAAACAUAUUAUCCAGACAUAGUUAUUAUACA 2458

Qy 1223 TGGACGACCTTCTAGCTGGCGAGCACCTGGAGATCGCCAGCACCCGCGCAAGATCGAG 1282
Db 2459 UGGAUUAUUGUAGUAGGACUAGUAGAAUAGGAGCAUAGAAACAAAUAUAGAGG 2518
Qy 1283 AGTGGCGAAGACACTGCTGGCTGGGCTTCACCAACCCCGCAGAAAGACACAGAGG 1342
Db 2519 AGCUGAGACAAUCUCUUGAGGUGGGAUUUACCAACACAGACAAAACAUAGAGAG 2578
Qy 1343 AGCCCCCTTCTGTGTGATGGCTACGAGTGCACCCGACCAAGTGGAGCTGCAGCCCA 1402
Db 2579 AACUCCAUUCCUUGUAGGUGUUAUAGAACUCCAUUAGUAAUAGGACAGAGACGCUA 2638
Qy 1403 TCGAGCTTCCCGGAGAGAGTGGACCGTGAACAGCATCCAGAAGCTGGTGGCGAAGC 1462
Db 2639 UAGUGUCGCCAGAAAAGACAGACUGACUGUAUAGACAAUAGUAGGGAU 2698
Qy 1463 TGAAGTGGCGAGCAGATCTACCCCGCATCAAGTGGCGCAGCTGTGCAAGTGTGTC 1522
Db 2699 UGAUUGGGCAAGUACAGAUUACCCAGGAAUUAAGAAAGCAUUAUAGUAAUCCUUA 2758
Qy 1523 GCGCGCAAGGCCCTGACCGACATCGTGCCTGACCGAGGAGCGAGCTGGAGCTGG 1582
Db 2759 GAGGAACCAAGCAGACUACAGAGAAUACCAUACAGAAAGACAGAGCUAGAACUG 2818
Qy 1583 CCGAGAACCGCGAGATCTGCGGAGCCGCTGCACGGCTGTACTAGCACCCAGCAAG 1642
Db 2819 CAGAAAACAGAGACAUUUAUAAAGAACACAGUACUAGGAGUUAUAGACCCCAUAAAG 2878
Qy 1643 ACCTGTGGCGAGATCCAGAGCAGGCGCAGCAGCTGAGTGGACCTTACAGATCTAC 1702
Db 2879 ACUUAUAGAGAAUUAACAGAGCAGGCGCAAGGCCAAUGGACAUUACAAUUAUAG 2938
Qy 1703 AGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCAGATCGCACCCGACCCACCAACG 1762
Db 2939 AGCCAUUUAUUAUUCUGAAACAGAAAUUAUAGCAUAGAGGUGGCCACACUUAUAG 2998
Qy 1763 ACCTGAGACGCTGACCGAGGCGCTGCAGAAAGATCGCCATGGAGAGCATCTGTATG 1822
Db 2999 AUGUAAACAAUUAACAGAGCAGUGCAAAAUAAUACACAGAAAGCAUAGUAAUAGG 3058
Qy 1823 GCAAGACCCCAAGTTCGGCTGCCATCCAGAGGAGACCTGGGAGACCTGGTGGACCG 1882
Db 3059 GAAAGACUCCUAAUUAUUAUACUACCAUAAAGAAACAAUAGGAAACAUUGGAGAG 3118
Qy 1883 ACTACTGCGAGCGCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCGCTGTGTA 1942
Db 3119 AGUUAUGGCAAGCCACUGGAUUCUGAGUGGAGUUAUUAUACCCUCCUUAUAGUA 3178
Qy 1943 AGCTGTGTACAGCTGGAGAGGAGCCATCATCTGGCGCGCAGACCTTCTACGTGGAG 2002
Db 3179 AAUUAUGGUACCAUAGAGAAAGAACCCUAGUAGGAGCAGAAACCUUUAUAGUAUG 3238
Qy 2003 GCGCGCCCAACCGGAGACCAAGATCGCAAGCGCGCTACGTGACCCAGCGGCGCGC 2062
Db 3239 GGGCAGCUAACAGGAGACUAAUUAAGAAAGCAGGAUUGUUAUACAAAGAAAGAGAC 3298
Qy 2063 AGAAGATCGTGGCTGACCGAGACCAACACAGAGACCGAGCTGCAGCGCATCCAGC 2122
Db 3299 AAAAGGUUUGCCUUAACUACACACAAAUACAGAAACUAGUUAUACAGCAUUUAU 3358
Qy 2123 TGGCCCTCAGGACAGCGGCGAGCGAGGTGAACATCGTGCAGCAGCAGCTACGCCCTGG 2182
Db 3359 UAGCUUUGCAGGAUUCAGGAUUAAGAAUAAACAUAGUAAACAGACUACAAUAGUAUG 3418
Qy 2183 GCATCATCCAGGCGCCCGCAGACAGCGAGCGAGCTGGTGAACAGATCATCTGAGC 2242
Db 3419 GAUUAUUCAGAGCAGACCAUAAUAAAGUAGAAUAGUUAUAGUUAUUAUUAUAGAG 3478
Qy 2243 AGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGCCCGCCACAGGCGCATCGCG 2302
Db 3479 AGUUAUAAAGGAAAGGAAAGGCUUAUUGGCAUGGAGUAGCAGCACACAAAGGAAUUGGAG 3538

LOCATION: (106)..(1641)
OTHER INFORMATION: encodes proteins of viral inner core
NAME/KEY: misc.feature
LOCATION: (1401)..(3617)
OTHER INFORMATION: encodes viral proteins but not integrase
NAME/KEY: misc.feature
LOCATION: (3708)..(5715)
OTHER INFORMATION: regulates high-level production of HIV genes
NAME/KEY: misc.feature
LOCATION: (3847)..(5944)
OTHER INFORMATION: encodes protein regulating the transfer of RNA to cytoplasm
NAME/KEY: misc.feature
LOCATION: (3939)..(4181)
OTHER INFORMATION: gene participates in viral assembly and budding
NAME/KEY: misc.feature
LOCATION: (4099)..(4941)
OTHER INFORMATION: encodes truncated form of viral coat protein
NAME/KEY: misc.feature
LOCATION: (6098)..(9918)
OTHER INFORMATION: vaccine vector pGAI

US-09-798-675-5

Query Match 43.5%; Score 1073.2; DB 10; Length 9918;
Best Local Similarity 67.2%; Pred. No. 1.9e-169;
Matches 1618; Conservative 0; Mismatches 738; Indels 52; Gaps 5;

QY	14	TGCGCCAGCCATGAGCCAGG---CCACACAGCGCCACATCCTCTGATGCGAGCGAGCAACT	70
DB	1193	TGCGTGAAGCAATGAGCCAAAGTAAACAAATACAGCTACCAATATGATGCGAGAGGCANNT	1252
QY	71	TCAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCGGCACATGCGCC	130
DB	1253	TTAGGACCAAGAAAGATGTTTAAGAGCTTCAATAGCGGCAAGAGGCGCACACAGCCA	1312
QY	131	GCACCTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTCCGCGAGGAGGCGGCACACAGA	190
DB	1313	GAATTTGAGGGCGCCCTAGGAAAGAGGCGAGCTGGAAAGCGGAAAGAGGAGGACACAAA	1372
QY	191	TGAAGGACTCACCGAGCGCGAGCCAACTCTTCGCGAGGAGACCTTGGCCCTTCCCGCAGG	250
DB	1373	TGAAGATTTGACTGAGAGACAGGCTAA-TTTTTTGGGAGAGATCTGGCCCTTCCCTACAG	1431
QY	251	GCAGGCGCCCGAGT-----TCCCGCAGCG	274
DB	1432	GGAAGCGCAGGAATTTCTTCAGAGCAGACAGACCAACAGCCCGCCACCATTTCTTCAG	1491
QY	275	AGCAGACCCGCGCACAGCCCGCCACAGCGCGAGCTGCGAGGTGCGCGG-----CGACA	328
DB	1492	AGCAGACCGAGCGCCACAGCCCGCCACAGAGAGAGCTTCAGGCTGCGGGGTAGAGACAA	1551
QY	329	ACCCCGCAGCGAGCGCGCGCGAGCGCCAGGCGCA-----CCCTGAACCTTCCCGCAGA	382
DB	1552	ACTCCCGCTCAGAAGCAGGAGCGGATAGACAGGAAGTATCTTTAACTTCCCTCAGA	1611
QY	383	TCACCTGTGGAGCGCCCGCTGGTGAAGATCAAGTGGCGGCGCCAGATCAAGAGCGCC	442
DB	1612	TCACCTTTTGGCAGCAGCCCTCGTCAATAAAGATAGGGGGGCAACTAAAGGAAGCTC	1671
QY	443	TGCTGGACCGCGCGCAGCACCGCTGCTGGAGGAGATGAGCTCGCGCGCAAGTGGGA	502
DB	1672	TATTGATACAGGAGCAGATGATACAGATATTAGAAAGATGAGTTTCCAGGAAGATGGA	1731
QY	503	AGCCCAAGATGATGGCGGCGATCGCGGCTTCATCAAGGTGCGCGCAGTAGACACAGATCC	562
DB	1732	AACCAAAATGATAGGGGAATTGGAGGTTTATCAAGATTAAGACAGATATGATCAGATAC	1791
QY	563	TGATCGAGATCTGGGCAAGAGCCATCGCGCCGCTGCTGATGCGGCGCCACCGCCGTGA	622
DB	1792	TCATAGAAATCTGTGGACATAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGCA	1851
QY	623	ACATCATCGCGCAACATGCTGACCCAGCTGGGCTGCACCCCTGAACCTTCCCGCATCAGC	682
DB	1852	ACATAATTGGAAGAAATCTGTGACTCAGATTTGGTTGCACCTTTAAATTTTCCCATTAGCC	1911

QY	683	CCATCGAGACCGTGCCTGTAAGCTGAAGCCCGCATGAGCGCCCAAGGTGAAGCAGT	742
DB	1912	CTATTGAGACTGTACCAGTAATAATTAAGCCAGGATGATGATGCCCAAAAGTTAAACAAT	1971
QY	743	GGCCCTGACCGAGGAGAGATCAAGCCCTGACCCCATCTCGGAGGAGATGAGGAAGG	802
DB	1972	GGCCATTGACAGAAGAAATAAAGCATTTAGTAGAAATTTGTACAGAAATGGAAGG	2031
QY	803	AGGGCAAGATCACCAGATCGGCCCGGAGAACCCCTACACACCCCGCTGTTCGCCATCA	862
DB	2032	AAGGGAAATTTCAAAATTTGGCCCTGAGAATCCATACATACTCCAGTATTTGCCATAA	2091
QY	863	AGAAGAAGACACACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAACAAGCGCA	922
DB	2092	AGAAAAGACAGTACTAATGGAAGAAATTTAGTAGATTTTCAGAGACTTTAATAAGAGAA	2151
QY	923	CCAGGACTTCTGGAGGTGCACTGGGCATGCCCGACCCCGCGGCTCGAAGAAGAGA	982
DB	2152	CTCAAGACTTCTGGGAAGTTCAATTTAGGAATAGCACATCCCGCAGGGTTAAAAAGAAA	2211
QY	983	AGAGCGTGACCGTCTGGAGCTGGCGGACGCTTTCAGCGTGCCTCGAGGAGACT	1042
DB	2212	AATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTCAGTTCCTTATAGTAGAAGCT	2271
QY	1043	TCGCAAGTACACCGCTTCCACCATCCCGAGCATCAACAACAGAGACCCCGCGCATCGCT	1102
DB	2272	TCAGGAGTATATGCAATTTACCATACCTAGTATAACAAATGAGACACCGAGGATTAGAT	2331
QY	1103	ACCAGTACAACTGTGTCGCCAGGCTGGAAGGAGCGCCAGCAGTCTTTCAGAGCAGCA	1162
DB	2332	ATCAGTACAAATGTCTTCCACAGGGATGGAAGGATCACCAGCAATATTCCAAGTAGCA	2391
QY	1163	TGACCAAGATCTGGAGCGCTTCCGCGCCGCAACCCCGAGATCGTGATCTACCAAGTACA	1222
DB	2392	TGACAAAATCTTAGAGCCCTTTAAAAACAAAATCCAGACATAGTTTATCTATCAATACA	2451
QY	1223	TGACCACTCTACGTGGGCGAGCACTTGAGATTCGCGCAGCAGCCGCGCAAGATCGAGG	1282
DB	2452	TGAACGATTTGTATGATAGGATCTGACTTAGAATAAGGCGAGCATAGAACAAAATAGAGG	2511
QY	1283	AGTGGCAGACCTGCTGCGCTGGGCTTACCACCCCGCAACAGAACCCAGAGG	1342
DB	2512	AGTGGACAACTCTCTTGGGTGGGACTTACCACCGAGCAAAAACATCAGAAAG	2571
QY	1343	AGCCCGCTTCTGTGGATGGGTACAGCTGCAACCCCGCAACAGTGGACCGTCAGGCCA	1402
DB	2572	AACTCCATTTCTGTGGTGGTTATGAACTCCATCTGATATAATGACAGTACAGCCTA	2631
QY	1403	TCAGCTGCGCGAGAGGAGCTGGACCGTGAACGACATCCAGAACGCTGGTGGGCAAGC	1462
DB	2632	TAGTGTGTCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAAGTTAGTGGGAAAT	2691
QY	1463	TGAACTGGGCGAGCAGATCTACCCCGCATCAAGGTGCGCGAGCTGTGAAGCTGTGC	1522
DB	2692	TGAATACCGAAGTCAATTTCCCGAGGATTAAGTAAGCAATTTATGTAACCTCTTA	2751
QY	1523	CGCGCGCAAGGCGCTGACCGACATCTGTCGCCCTTGACCGAGGAGCGGAGCTGGAGTGG	1582
DB	2752	GAGCAACCAAGCACTAACAGAAGTAATACCACTAACAGAAGAAAGCAGAGCTAGAACTGG	2811
QY	1583	CCGAGAACCGGAGATCTTGGCGAGCGGTGACGGGTGTACTACGACCCCGAGCAAGG	1642
DB	2812	CAGAAAAACAGAGATTTCTAAAAGAACCGTACATGAGTGTATTATGACCCATCAAAAG	2871
QY	1643	ACCTGGTGGCGGAGATCCAGAAGCAGGCGCACAGCTGGAGCTTACCAGATCTACCAAG	1702
DB	2872	ACTTAATAGCAGAAATACAGAGCGGCGGAGCCCAATGACATATCAAAATTTATCAAG	2931
QY	1703	AGCCCTTCAAGAACTGAAGACCGGCAAGTACGCCAAGATCGCACCGCCGACCAACAGC	1762
DB	2932	AGCCATTTAAAAATCTGAAAACACAGAAAAATATGCAAGAATGAGGGGTGCCCCACACTAATG	2991

QY	1763	ACGTGAAGCAGCTGACCGAGCGCGTGCAGAGATCGCATGGAGAGACTGCTGATCTGGG	1822
Db	2992	ATGTAAACAATTAACAGAGCGAGTGCAAAAAATAACACAGAAGCATAGTAATATGGG	3051
QY	1823	GCAAGACCCCAAGTTCCGCTGCCCATCCAGAAGGAGACCTGGGAGACCTGGTGACCG	1882
Db	3052	GAAGACTCTAAATTTAACTACCCTATACAAGAGAAACATGGAAACATGGTGACAG	3111
QY	1883	ACTACTGGCAGCCACCTGGATCCCCGAGTGGGAGTTCTGTGAACACACCCGCCCTCGTGA	1942
Db	3112	AGTATTGGCAAGCCACCTGGATTCTCTGAGTGGGAGTTTGTTAATACCCCTCTTAGTGA	3171
QY	1943	AGCTGTGTTACAGCTGAGAGAGAGCCCATCATCGGCGCGGAGACTTCTACGTGGACG	2002
Db	3172	AATATTGGTACCAGTTGAGAGAAAGAACCCCATAGTAGGAGCAGAAACCTTCTATGTAGATG	3231
QY	2003	GGCGCCCAACCGCAGACCAAGATCGGCAAGCGCGCTACGTGACCGACCGGGCGCGC	2062
Db	3232	GGCAGCTAACAGGAGACTAAATTAGSANAAGCAGGATATGTTACTACAAAGGAAGAC	3291
QY	2063	AGAAGATCGTGAGCTGACCGAGACACCAACACAGAAGCCGAGCTGCAGGGCCATCCAGC	2122
Db	3292	AAAAGGTTGCCCCCTAACTAACACAACTACGAAACTCAGTTACAAGCAATTTATC	3351
QY	2123	TGGCCCTGCAGGACAGCGCAGGTGACATCTGTGACCGACAGCCAGTACGCCCTGG	2182
Db	3352	TAGCTTTGCAGGATTCAGGATTAGAAGTAAACATAGTAACAGACTCACATATGCATTAG	3411
QY	2183	GCATCATCCAGCCCGCCGACAGAGCGAGCGAGCTGGTGAACAGACTCATCGAGC	2242
Db	3412	GAATCATTCAGCAACACCATAAAGTGAATCAGAGTTAGTCAATCAATAATAGACG	3471
QY	2243	AGCTGATCAAGAGAGAGGTGTACTGTAGCTGGGTGCCCGCCCAAGGGCATCGGGC	2302
Db	3472	AGTTAATAAAAAGGAAAGTCTTACTGGCATGGGTACCAGCACACAAAGGAATTTGGAG	3531
QY	2303	GCAACGAGCAGATGCACAAGCTGTGTAGCAAGGSCATCCGACAGGTGCTGTTCCCTGGAGC	2362
Db	3532	GAATGACACAGTAGATAAATTAGTCACTGCTGGAATFCAGGAANAATCTATTTTAGATG	3591
QY	2363	GCATCGAT	2370
Db	3592	GAATAGAT	3599

Search completed: February 10, 2003, 21:10:57
job time : 153.929 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:53:38 ; Search time 6249 Seconds
(without alignments)
11470.669 Million cell updates/sec
Title: US-09-610-313-31
Perfect score: 2463
Sequence: 1 gtcagccacccatggccga.....gggctagcaccggtaattc 2463

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_on.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2463	100.0	2463	6	AX455915 Sequence
2	2442.2	99.2	2469	6	AX455914 Sequence
3	2436.2	98.9	2457	6	AX455916 Sequence
4	2027.6	82.3	9166	6	AX427930 Sequence
5	2019.2	82.0	9788	6	AX427936 Sequence
6	2005.2	81.4	9169	6	AX427931 Sequence
7	2001.8	81.3	9194	6	AX427926 Sequence
8	2000.2	81.2	9194	6	AX427925 Sequence
9	2000.2	81.2	12411	6	AX427927 Sequence
10	1996.8	81.1	9785	6	AX427938 Sequence
11	1993.4	80.9	9170	6	AX427928 Sequence
12	1993.4	80.9	9782	6	AX427935 Sequence
13	1993.4	80.9	9783	6	AX427934 Sequence
14	1993.4	80.9	9792	6	AX427932 Sequence
15	1991.8	80.9	9189	6	AX427921 Sequence
16	1988.8	80.7	4353	12	AF287353 Synthetic
17	1972.2	80.1	3015	6	AX455946 Sequence
18	1955.6	79.4	3009	6	AX455987 Sequence
19	1932	78.4	4352	12	AF287352 Synthetic
20	1886.6	76.6	2577	6	AX457088 Sequence
21	1879	76.3	4343	6	AX306428 Sequence
22	1875	76.1	4341	6	AX306429 Sequence
23	1865.4	75.7	1977	6	AX455954 Sequence
24	1857.8	75.4	1965	6	AX455952 Sequence
25	1852.4	75.2	7897	6	AX427923 Sequence
26	1798.8	73.0	3024	12	AF287355 Synthetic
27	1784	72.4	9407	6	AX427937 Sequence
28	1651.8	67.1	4307	6	AX019132 Sequence
29	1632.6	66.3	4307	6	AX035453 Sequence
30	1632.6	66.3	4307	6	AX056836 Sequence
31	1624.6	66.0	4307	6	AX283597 Sequence
32	1624.6	66.0	4307	6	AX283607 Sequence
33	1624.6	66.0	4327	6	AX035465 Sequence
34	1624.6	66.0	4327	6	AX035463 Sequence
35	1624.6	66.0	4642	6	AX035464 Sequence
36	1624.6	66.0	4642	6	AX035463 Sequence
37	1624.6	66.0	9772	6	AX283584 Sequence
38	1581.2	64.2	1680	6	AX455960 Sequence
39	1575.2	64.0	1668	6	AX455962 Sequence
40	1566	63.6	8366	6	AX180752 Sequence
41	1530.6	62.1	4338	6	AX180747 Sequence
42	1518.2	61.6	3981	6	AX306430 Sequence
43	1300.8	52.8	4288	6	AX149648 Sequence
44	1255.8	51.0	9010	14	AF110974
45	1255.8	51.0	9010	14	AF110975

ALIGNMENTS

RESULT 1
AX455915
LOCUS AX455915 2463 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 31 from Patent WO0204493.
ACCESSION AX455915
VERSION AX455915.1 GI:21714900
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS zur Megede,J., Barnett,S.W., Engelbrecht,S. and van Rensburg,E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 31 17-JAN-2002;

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

Location/Qualifiers
1. .2463
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975YM"

BASE COUNT 567 a 835 c 759 g 302 t

ORIGIN

Query Match 100.0%; Score 2463; DB 6; Length 2463;

Best Local Similarity 100.0%; Pred. No. 1.9e-249;

Matches 2463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACGCCACCATGGCCGAGGCATGAGCCAGCCACCGAGCGCCAACTCCTGATSCAG 60
DB 1 GTGACGCCACCATGGCCGAGGCATGAGCCAGCCACCGAGCGCCAACTCCTGATSCAG 60
QY 61 CGCAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTGTTCCTCAACTCGCGCAAGAGGGC 120
DB 61 CGCAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTGTTCCTCAACTCGCGCAAGAGGGC 120
QY 121 CATATCGCCGCAACTCGCGGCCGCCGCCGCCGCAAGAGGGCTGTGGAAGTGGGGCAAGGAG 180
DB 121 CATATCGCCGCAACTCGCGGCCGCCGCCGCCGCAAGAGGGCTGTGGAAGTGGGGCAAGGAG 180
QY 181 GGCACACAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGCACACAGATGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTCCGCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGAGCAACCGCGCAACAGAGCCGACC 300
DB 241 TTCCGCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGAGCAACCGCGCAACAGAGCCGACC 300
QY 301 AGCGCGAGCTGACAGTGGCGGGGCAACCCCGCAGCGAGCGCCGCCGCGAGCGCCAG 360
DB 301 AGCGCGAGCTGACAGTGGCGGGGCAACCCCGCAGCGAGCGCCGCCGCGAGCGCCAG 360
QY 361 GGCACCTGAACTTCCCGCCAGATCACCTGTGGCAGCGCCCGCTTGTGAGCATCAAGGTG 420
DB 361 GGCACCTGAACTTCCCGCCAGATCACCTGTGGCAGCGCCCGCTTGTGAGCATCAAGGTG 420
QY 421 GCGCGCAGATCAAGAGGCGCCTGCTGACACCGCGCGCAGACACCGTGTGAGGAG 480
DB 421 GCGCGCAGATCAAGAGGCGCCTGCTGACACCGCGCGCAGACACCGTGTGAGGAG 480
QY 481 ATGAGCTGCCGCGCAAGTGGAGCCCAAGATGATCGCGCATCGCGGCTTCATCAAG 540
DB 481 ATGAGCTGCCGCGCAAGTGGAGCCCAAGATGATCGCGCATCGCGGCTTCATCAAG 540
QY 541 GTGCGCCAGTACGACCATGCTGATCGAGATCTGCGGCAAGAGAGGCGCATCGGACCGTG 600
DB 541 GTGCGCCAGTACGACCATGCTGATCGAGATCTGCGGCAAGAGAGGCGCATCGGACCGTG 600
QY 601 CTGATGCGGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGTGC 660
DB 601 CTGATGCGGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGTGC 660
QY 661 ACCCTGAACTTCCCGCATCAGCCCGCATCGAGACCGTGGCGGTGAAGCTGAAGCCCGCATG 720
DB 661 ACCCTGAACTTCCCGCATCAGCCCGCATCGAGACCGTGGCGGTGAAGCTGAAGCCCGCATG 720
QY 721 GACGGCCCAAGGTGAAGAGTGGCCCGTACCAGGAGAGAGATCAAGGGCCCTGACCGCC 780
DB 721 GACGGCCCAAGGTGAAGAGTGGCCCGTACCAGGAGAGAGATCAAGGGCCCTGACCGCC 780
QY 781 ATCTCGGAGGAGATGGAGAGGGCAAGATCACCAAGATCGCGCCCGCGCAAGCCCTTAC 840
DB 781 ATCTCGGAGGAGATGGAGAGGGCAAGATCACCAAGATCGCGCCCGCGCAAGCCCTTAC 840
QY 841 AACACCCCGGTGTCGGCATCAAGAGAGAGGACACCAAGTGGCGGCAAGCTGTGAC 900
DB 841 AACACCCCGGTGTCGGCATCAAGAGAGAGGACACCAAGTGGCGGCAAGCTGTGAC 900

QY 901 TTCCGGAGCTGAACAGCGACCCAGGACTTCTGGAGGTGAGCTGGCATCCCCAC 960
DB 901 TTCCGGAGCTGAACAGCGACCCAGGACTTCTGGAGGTGAGCTGGCATCCCCAC 960
QY 961 CCCCGCGGCTGAAGAGAGAGAGAGCGTACCGTGTGACGTGGCGGCGGCTACTTC 1020
DB 961 CCCCGCGGCTGAAGAGAGAGAGAGCGTACCGTGTGACGTGGCGGCGGCTACTTC 1020
QY 1021 AGCGTCCCGCTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCAAC 1080
DB 1021 AGCGTCCCGCTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCAAC 1080
QY 1081 AAGCAGACCCCGGCTACCGCTACAGTACAAAGTGTGCCCCAGGGCTGGAAGGGCAGC 1140
DB 1081 AAGCAGACCCCGGCTACCGCTACAGTACAAAGTGTGCCCCAGGGCTGGAAGGGCAGC 1140
QY 1141 CCAGCATCTTCCAGAGCAGATGACCAAGATCCTGAGGCGCTTCGCGGCCGCAAGCCC 1200
DB 1141 CCAGCATCTTCCAGAGCAGATGACCAAGATCCTGAGGCGCTTCGCGGCCGCAAGCCC 1200
QY 1201 GAGATCGTGTATCTACCAAGGCCCTTCTAGTGGCGAGCGACCTGGAGATCGGCGCAGCAC 1260
DB 1201 GAGATCGTGTATCTACCAAGGCCCTTCTAGTGGCGAGCGACCTGGAGATCGGCGCAGCAC 1260
QY 1261 CGCGCCAAAGATCGAGGAGCTGCGCAAGCACTGCTGCGCTGGGGCTTACACACCCCGAC 1320
DB 1261 CGCGCCAAAGATCGAGGAGCTGCGCAAGCACTGCTGCGCTGGGGCTTACACACCCCGAC 1320
QY 1321 AAGAAGCACCAGAGAGGAGCGCCCTTCTGTGATGGCTACGAGCTGCACCCCGACAG 1380
DB 1321 AAGAAGCACCAGAGAGGAGCGCCCTTCTGTGATGGCTACGAGCTGCACCCCGACAG 1380
QY 1381 TGGACCTGTGACGCCATCTGAGCTGCCGAGAGAGAGCTGGACCGTGAACGACATCCAG 1440
DB 1381 TGGACCTGTGACGCCATCTGAGCTGCCGAGAGAGAGCTGGACCGTGAACGACATCCAG 1440
QY 1441 AAGTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGGCGATCAAGGTGGCGCAG 1500
DB 1441 AAGTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGGCGATCAAGGTGGCGCAG 1500
QY 1501 CTGTGCAAGCTGTGCGGGCGCAAGGCCCTGACGACATCTGCGCCCTGACCGAGGAG 1560
DB 1501 CTGTGCAAGCTGTGCGGGCGCAAGGCCCTGACGACATCTGCGCCCTGACCGAGGAG 1560
QY 1561 GCGCAGCTGGAAGCTGGCGGAGAACCGGAGATCTTGGCGGAGCGCCGTGACCGGCTGTAC 1620
DB 1561 GCGCAGCTGGAAGCTGGCGGAGAACCGGAGATCTTGGCGGAGCGCCGTGACCGGCTGTAC 1620
QY 1621 TACGACCCCAAGAGGAGCTGGTGGCGGAGATCAGAGAGGCGGCGACCGAGTGGACC 1680
DB 1621 TACGACCCCAAGAGGAGCTGGTGGCGGAGATCAGAGAGGCGGCGACCGAGTGGACC 1680
QY 1681 TACGAGATCTACGAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGGCC 1740
DB 1681 TACGAGATCTACGAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGGCC 1740
QY 1741 ACCGCCACACCAAGCAGGTGAAGCAGTACCGGCGGCTGCGAAGAGATCGCCATGAG 1800
DB 1741 ACCGCCACACCAAGCAGGTGAAGCAGTACCGGCGGCTGCGAAGAGATCGCCATGAG 1800
QY 1801 AGCATGTGTGCTGGGGCAAGACCCCAAGTTCGCGCTTCCCATCCAGAGGAGACCTGG 1860
DB 1801 AGCATGTGTGCTGGGGCAAGACCCCAAGTTCGCGCTTCCCATCCAGAGGAGACCTGG 1860
QY 1861 GAGACCTGTGTGAGCGGCTACTGCGAGCGCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
DB 1861 GAGACCTGTGTGAGCGGCTACTGCGAGCGCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
QY 1921 ACCGCCCGCTGTGTAAGCTGTGTACAGCTGGAGAGAGGCGCCATCATCGGCCCGAG 1980
DB 1921 ACCGCCCGCTGTGTAAGCTGTGTACAGCTGGAGAGAGGCGCCATCATCGGCCCGAG 1980
QY 1981 ACCTTCTACGTGGAGCGGCGGCCCAAGCCGAGACCAAGATCGGCAAGCGCGCTACGTG 2040

Db 1981 ACCTTACCTGGACGGCGCCGCAACGGGAGACCAAGATCGCAAGGCGGCTACGTG 2040
QY 2041 ACCGACGGGCGGCGGAGAGATCGTGAGCTGACCGGAGACCAACCAAGACCGAG 2100
Db 2041 ACCGACGGGCGGCGGAGAGATCGTGAGCTGACCGGAGACCAACCAAGACCGAG 2100
QY 2101 CTGAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGAGGTGAACATCGTGACCGAC 2160
Db 2101 CTGAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGAGGTGAACATCGTGACCGAC 2160
QY 2161 AGCCAGTACGCCCTGGGATCATCTCAGGCCAGCCGACAGAGCGAGAGGAGCTGGTG 2220
Db 2161 AGCCAGTACGCCCTGGGATCATCTCAGGCCAGCCGACAGAGCGAGAGGAGCTGGTG 2220
QY 2221 AACAGATCATCGAGAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGCCCGCC 2280
Db 2221 AACAGATCATCGAGAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGCCCGCC 2280
QY 2281 CACAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGTGTAGCAAGGGCATCCGCAAG 2340
Db 2281 CACAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGTGTAGCAAGGGCATCCGCAAG 2340
QY 2341 GTGCTGTCTTGGACGGATCGATGGCGGATCGATCGATCTACAGTACAGGAGCAGCTG 2400
Db 2341 GTGCTGTCTTGGACGGATCGATGGCGGATCGATCGATCTACAGTACAGGAGCAGCTG 2400
QY 2401 TACGTGGGACGGCGCCCTAGGATCGATTAAAGCTTCCGGGGCTAGCACCGGTGAA 2460
Db 2401 TACGTGGGACGGCGCCCTAGGATCGATTAAAGCTTCCGGGGCTAGCACCGGTGAA 2460
QY 2461 TTC 2463
Db 2461 TTC 2463

RESULT 2
AX455914
LOCUS AX455914 2469 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 30 from Patent WO0204493.
ACCESSION AX455914
VERSION AX455914.1 GI:21714899

SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

JOURNAL
zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
Patent: WO 0204493-A 30 17-JAN-2002;

CHIRON CORPORATION (US); University of Stellenbosch (ZA)
Location/Qualifiers
1. .2469

FEATURES
source
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975(+)"

BASE COUNT 571 a 833 c 761 g 304 t
ORIGIN

Query Match 99.2%; Score 2442.2; DB 6; Length 2469;
Best Local Similarity 99.6%; Pred. No. 2.9e-247;
Matches 2460; Conservative 0; Mismatches 3; Indels 6; Caps 1;

QY 1 GTCAGCCACCATGGCCGAGGCGCATGAGCCAGCCAGCCAGCCACATCTGATCGAG 60
Db 1 GTCAGCCACCATGGCCGAGGCGCATGAGCCAGCCAGCCAGCCACATCTGATCGAG 60

QY 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTGTCAACTCGCGCAAGGAGGGC 120
Db 61 CGCAGCAACTTCAAGGCGCCCAAGCGCATCATCAAGTGTCAACTCGCGCAAGGAGGGC 120

QY 121 CACATCGCCGCAACTGCGCGGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGAAGGAG 180

Db 121 CACATCGCCGCAACTGCGCGGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGAAGGAG 180
QY 181 GGCACACAGATGAGGAGCTGACCGAGCGCCAGGCGCAACTTCTTCGCGGAGGAGACTGGCC 240
Db 181 GGCACACAGATGAGGAGCTGACCGAGCGCCAGGCGCAACTTCTTCGCGGAGGAGACTGGCC 240
QY 241 TTCCCGCCAGGCGCAAGGCGCGGAGTTCCCGAGGAGCAAGCCGCGCCCAAGCGCCCAACC 300
Db 241 TTCCCGCCAGGCGCAAGGCGCGGAGTTCCCGAGGAGCAAGCCGCGCCCAAGCGCCCAACC 300
QY 301 AGCGCGAGCTGCGAGGTGCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCAG 360
Db 301 AGCGCGAGCTGCGAGGTGCGCGGCGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCAG 360
QY 361 GGCACCTGAACTTCCCGCCAGATCACCTGATCGAGATCTGCGGCAAGGCGCTTCAATCAAG 420
Db 361 GGCACCTGAACTTCCCGCCAGATCACCTGATCGAGATCTGCGGCAAGGCGCTTCAATCAAG 420
QY 421 GGCAGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCGGAG 480
Db 421 GGCAGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCGGAG 480
QY 481 ATGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGCATCGCGGCGCTTCAATCAAG 540
Db 481 ATGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGCGGCGCATCGCGGCGCTTCAATCAAG 540
QY 541 GTGCGCAGTACGACCGATCTGATCGAGATCTGCGGCAAGGCGCGCGCGCGCGCGCGCG 600
Db 541 GTGCGCAGTACGACCGATCTGATCGAGATCTGCGGCAAGGCGCGCGCGCGCGCGCGCG 600
QY 601 CTGATCG 660
Db 601 CTGATCG 660
QY 661 ACCCTGAACCTTCCCGCCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 661 ACCCTGAACCTTCCCGCCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GACG 780
Db 721 GACG 780
QY 781 ATCTGCGAGGAGTGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db 781 ATCTGCGAGGAGTGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 AACACCCCGGTGTCGCCATCAAGAAGGAGCGACAGCAAGTGGCGCGCAAGCTGGTGAC 900
Db 841 AACACCCCGGTGTCGCCATCAAGAAGGAGCGACAGCAAGTGGCGCGCAAGCTGGTGAC 900
QY 901 TTCCGCGAGCTGAACAGCG 960
Db 901 TTCCGCGAGCTGAACAGCG 960
QY 961 CCCCG 1020
Db 961 CCCCG 1020
QY 1021 AGCGTGCCTTGGACGAGGAGTTCGCCAAGTACACCGCGCTTACCATCTCCCGAGCATCAAC 1080
Db 1021 AGCGTGCCTTGGACGAGGAGTTCGCCAAGTACACCGCGCTTACCATCTCCCGAGCATCAAC 1080
QY 1081 AACGAGACCCCG 1140
Db 1081 AACGAGACCCCG 1140
QY 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTGAGAGCGCTTCCCGCGCGCGCGCGCG 1200
Db 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTGAGAGCGCTTCCCGCGCGCGCGCGCG 1200
QY 1201 GAGATCGTGTACTACCA-----GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254
Db 1201 GAGATCGTGTACTACCA-----GGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1254

Db 1201 GAGATCGTGATCTACAGTACATGAGACCTGTAGTGGGAGGACCTGGAGATCGGC 1260
QY 1255 CAGCACCGCGCCAAAGTCAGGAGCTGCGCAGACCTCTGCTGGCTGGGCTTCACCAAC 1314
Db 1261 CAGCACCGCGCCAAAGTCAGGAGCTGCGCAGACCTCTGCTGGCTGGGCTTCACCAAC 1320
QY 1315 CCGCACAAGAAGCACCAGAGGAGCCCTTCTGCTGGATGGCTTACAGCTGCAACCC 1374
Db 1321 CCGCACAAGAAGCACCAGAGGAGCCCTTCTGCTGGATGGCTTACAGCTGCAACCC 1380
QY 1375 GACAACTGAGCCGTGAGCCCATCGAGCTGCGGAGAGAGAGCTGGACCGTGAACGAC 1434
Db 1381 GACAACTGAGCCGTGAGCCCATCGAGCTGCGGAGAGAGAGCTGGACCGTGAACGAC 1440
QY 1435 ATCCAGAACTGCTGGGCAAGCTGAAGTGGGCGAGCCAGATCTACCCCGGCATCAAGGTG 1494
Db 1441 ATCCAGAACTGCTGGGCAAGCTGAAGTGGGCGAGCCAGATCTACCCCGGCATCAAGGTG 1500
QY 1495 CGCAGCTGTGAAGCTGCTGGCGCGGCCAAGGCCCTGACCCACATCTGTCCTCTGACC 1554
Db 1501 CGCAGCTGTGAAGCTGCTGGCGCGGCCAAGGCCCTGACCCACATCTGTCCTCTGACC 1560
QY 1555 GAGGAGCGGAGCTGAGCTGCGCGGAGAACCGGAGATCTGCGGAGCCCGTGCACGGC 1614
Db 1561 GAGGAGCGGAGCTGAGCTGCGCGGAGAACCGGAGATCTGCGGAGCCCGTGCACGGC 1620
QY 1615 GTGTACTAGACCCAGCAGGAGCTGCTGGCGGAGAACCGGAGATCTGCGGAGCCCGTGCACGGC 1674
Db 1621 GTGTACTAGACCCAGCAGGAGCTGCTGGCGGAGAACCGGAGATCTGCGGAGCCCGTGCACGGC 1680
QY 1675 TGGACCTACAGATCTACAGGAGCCCTTCAAGAACTGAAAGCCCGGAGTACGCGCAAG 1734
Db 1681 TGGACCTACAGATCTACAGGAGCCCTTCAAGAACTGAAAGCCCGGAGTACGCGCAAG 1740
QY 1735 ATGCGCAGCCGCCACACCAACGAGCTGAGAGCTGACCGAGCTGACCGAGCCCGTGCAGAAATCGCC 1794
Db 1741 ATGCGCAGCCGCCACACCAACGAGCTGAGAGCTGACCGAGCTGACCGAGCCCGTGCAGAAATCGCC 1800
QY 1795 ATGCGCAGATCTGATCTGGGCAAGACCCCAAGTTCGCTGCGCCATCCAGAGGAG 1854
Db 1801 ATGCGCAGATCTGATCTGGGCAAGACCCCAAGTTCGCTGCGCCATCCAGAGGAG 1860
QY 1855 ACCTGGAGACCTGCTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC 1914
Db 1861 ACCTGGAGACCTGCTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTC 1920
QY 1915 GTGAACACCCCGCTGCTGAGCTGTGTACCACTGAGAGAGAGCCCATCATCGGC 1974
Db 1921 GTGAACACCCCGCTGCTGAGCTGTGTACCACTGAGAGAGAGCCCATCATCGGC 1980
QY 1975 GCGGAGACCTTCTACGTGGACGCGCGCCCAAGCGGAGACCAAGATCGGCAAGGCGGC 2034
Db 1981 GCGGAGACCTTCTACGTGGACGCGCGCCCAAGCGGAGACCAAGATCGGCAAGGCGGC 2040
QY 2035 TAGCTGACCGACCGGCGCGGAGAGATCTGTGAGCTGACCGAGACCAACCAAGAG 2094
Db 2041 TAGCTGACCGACCGGCGCGGAGAGATCTGTGAGCTGACCGAGACCAACCAAGAG 2100
QY 2095 ACCGAGCTGAGGCGCATTCAGCTGGCCCTGAGGAGAGCGGAGAGAGTGAACATCGTG 2154
Db 2101 ACCGAGCTGAGGCGCATTCAGCTGGCCCTGAGGAGAGCGGAGAGAGTGAACATCGTG 2160
QY 2155 ACCGAGAGCAGTACGCTTGGGCAATCTCCAGGCGCCAGCCGAGAGAGGAGCGAG 2214
Db 2161 ACCGAGAGCAGTACGCTTGGGCAATCTCCAGGCGCCAGCCGAGAGAGGAGCGAG 2220
QY 2215 CTGGTGAACCAAGATCTAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGGTG 2274
Db 2221 CTGGTGAACCAAGATCTAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGGTG 2280
QY 2275 CCGGCGCCACAGGGCATCGGCGGACACGAGAGATGACAAAGCTGTGTGAGCAAGGCGATC 2334
Db 2281 CCGGCGCCACAGGGCATCGGCGGACACGAGAGATGACAAAGCTGTGTGAGCAAGGCGATC 2340

QY 2335 CGCAAGTGTCTTCTGGACGGCATCGTGGCGGATCTACCACTACATGGAC 2394
Db 2341 CGCAAGTGTCTTCTGGACGGCATCGTGGCGGATCTACCACTACATGGAC 2400
QY 2395 GACCTGTAGTGGGAGCGGCGGCTTAGGATCGATTAAAGCTTCCCGGGGTAGCAC 2454
Db 2401 GACCTGTAGTGGGAGCGGCGGCTTAGGATCGATTAAAGCTTCCCGGGGTAGCAC 2460
QY 2455 GGTGAATTC 2463
Db 2461 GGTGAATTC 2469
RESULT 3
AX455916
LOCUS
DEFINITION Sequence 32 from Patent WO0204493.
ACCESSION AX455916
VERSION AX455916.1 GI:21714901
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 artificial sequences.
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
JOURNAL Patent: WO 0204493-A 32 17-JAN-2002;
CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
1..2457
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975YMMW"
BASE COUNT 566 a 837 c 754 g 300 t
ORIGIN
Query Match 98.9%; Score 2436.2; DB 6; Length 2457;
Best Local Similarity 99.6%; Pred. No. 1.2e-246;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;
QY 1 GTGCGCGCCACCATGTGCGGAGGCGCATGAGCGCCAGCCACAGCGCCCAACATCTCTGATGCAG 60
Db 1 GTGCGCGCCACCATGTGCGGAGGCGCATGAGCGCCAGCCACAGCGCCCAACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGAGGGC 120
Db 61 CGCAGCAACTTCAAGGGCGCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGAGGGC 120
QY 121 CACATCGCGCGCAACTGCGCGCGCCCGCCCGCAAGAGGGTCTGGAAGTGGCGCAAGGAG 180
Db 121 CACATCGCGCGCAACTGCGCGCGCCCGCCCGCAAGAGGGTCTGGAAGTGGCGCAAGGAG 180
QY 181 GGCACACAGATGAAGACTGCAACCGAGCGCCAGCGCAACTTCTTCCGCGAGGACCTGGCC 240
Db 181 GGCACACAGATGAAGACTGCAACCGAGCGCCAGCGCAACTTCTTCCGCGAGGACCTGGCC 240
QY 241 TTCCCGCCAGGCGCGCGCGAGTTCCCGAGCGAGCAACCGCGCCCAACAGCCCGCCACC 300
Db 241 TTCCCGCCAGGCGCGCGCGAGTTCCCGAGCGAGCAACCGCGCCCAACAGCCCGCCACC 300
QY 301 AGCGCGAGCTGCAAGTGGCGCGGAGACACCCCGCGAGGCGCGCGCGCGCGCGCGAG 360
Db 301 AGCGCGAGCTGCAAGTGGCGCGGAGACACCCCGCGAGGCGCGCGCGCGCGCGCGAG 360
QY 361 GGCACCTTGAACCTTCCCGCAGATCAACCTGTGGCAGCGCCCTTGGTGGAGCATCAAGGTG 420
Db 361 GGCACCTTGAACCTTCCCGCAGATCAACCTGTGGCAGCGCCCTTGGTGGAGCATCAAGGTG 420
QY 421 GCGGCGCAGATCAAGGAGGCGCTGTGGAGACCGCGCGCGCGCGCGCGCGCGCGCGAG 480
Db 421 GCGGCGCAGATCAAGGAGGCGCTGTGGAGACCGCGCGCGCGCGCGCGCGCGCGCGAG 480

QY 481 ATGAGCCTGCCGGCAAGTGGAGCCCAAGATGATCGCGGCATCGGCGCTTCATCAAG 540
DB 481 ATGAGCCTGCCGGCAAGTGGAGCCCAAGATGATCGCGGCATCGGCGCTTCATCAAG 540
QY 541 GTGCGCAGTACGACAGATCCCTGATCGAGATCTGCGGCAAGAGGCAATCGGCACCGTG 600
DB 541 GTGCGCAGTACGACAGATCCCTGATCGAGATCTGCGGCAAGAGGCAATCGGCACCGTG 600
QY 601 CTGATCGGCCCAACCCCGTGAACATCATCGGCCGGAACATGCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCCCAACCCCGTGAACATCATCGGCCGGAACATGCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAACCTTCCCATCATGACGCCCATCGAGACGCTGCCGTGAAGCTGAAGCCCGCATG 720
DB 661 ACCCTGAACCTTCCCATCATGACGCCCATCGAGACGCTGCCGTGAAGCTGAAGCCCGCATG 720
QY 721 GACGGCCCCAAGTGAAGCATGAGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
DB 721 GACGGCCCCAAGTGAAGCATGAGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGGAGAGATGAGAGAGGAGGCAAGATCAACAGATCGGCCCGCCGAGACCCCTAC 840
DB 781 ATCTGGAGAGATGAGAGAGGAGGCAAGATCAACAGATCGGCCCGCCGAGACCCCTAC 840
QY 841 AACACCCCGCTGTCGCCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 AACACCCCGCTGTCGCCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 TTCGGGAGCTGAACAGCGACCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCGCCAC 960
DB 901 TTCGGGAGCTGAACAGCGACCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCGCCAC 960
QY 961 CCGCGCGCTGAAG 1020
DB 961 CCGCGCGCTGAAG 1020
QY 1021 AGCGTGCCCTGGAGAGAGACTTCCGCAAGTACACGCCCTTACCATCCCGAGCATCAAC 1080
DB 1021 AGCGTGCCCTGGAGAGAGACTTCCGCAAGTACACGCCCTTACCATCCCGAGCATCAAC 1080
QY 1081 AAGAGACCCCGGCATCCGCTACCAAGTACAGTGTGTGCCCGCAGGCTGGAAGGGAGC 1140
DB 1081 AAGAGACCCCGGCATCCGCTACCAAGTACAGTGTGTGCCCGCAGGCTGGAAGGGAGC 1140
QY 1141 CCAGAGATCTTCAGAGCAGATGACCAAGATCTTGAGAGCCCTTCCGCGCCGGAACCCC 1200
DB 1141 CCAGAGATCTTCAGAGCAGATGACCAAGATCTTGAGAGCCCTTCCGCGCCGGAACCCC 1200
QY 1201 GAGATCGTGATACCAAGGCCCTTGTAGTGGGAGGAGCTGGAGATCGGCGCAGCAG 1260
DB 1201 GAGATCGTGATACCAAGGCCCTTGTAGTGGGAGGAGCTGGAGATCGGCGCAGCAG 1260
QY 1261 CGGCGCAAGATCGAGGAGCTGCGCAAGCCTGCTGCCTGGGCTTCCACACCCCGCAG 1320
DB 1261 CGGCGCAAGATCGAGGAGCTGCGCAAGCCTGCTGCCTGGGCTTCCACACCCCGCAG 1320
QY 1321 AAGAGCAGCAG 1380
DB 1321 AAGAGCAGCAG 1380
QY 1381 TGGACCTGAGCCCATCGAGCTGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1381 TGGACCTGAGCCCATCGAGCTGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 AAGCTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGGCATCAAGGTGCGCCAG 1500
DB 1441 AAGCTGTGGGCAAGCTGAATGGGCGCAGCAGATCTACCCCGGCATCAAGGTGCGCCAG 1500
QY 1501 CTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCGACATCTGTCGCCCTGACCGAGGAG 1560
DB 1501 CTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCGACATCTGTCGCCCTGACCGAGGAG 1560
QY 1561 GCGAGCTGAGCTGCGCGGAGAACCGCGAGATCTGCGCGAGCCCTGACGCGCTGTAC 1620
DB 1561 GCGAGCTGAGCTGCGCGGAGAACCGCGAGATCTGCGCGAGCCCTGACGCGCTGTAC 1620
QY 1621 TAGGACCCCAAGAGACCTGCTGGCGGAGATCCAGAAGCAGGCGCCACGACAGTGGACC 1680
DB 1621 TAGGACCCCAAGAGACCTGCTGGCGGAGATCCAGAAGCAGGCGCCACGACAGTGGACC 1680
QY 1681 TACAGATCTTACAGGAGCCCTTCAAGAACCTGAGACCCGCGCAAGTACGCAAGATCGGC 1740
DB 1681 TACAGATCTTACAGGAGCCCTTCAAGAACCTGAGACCCGCGCAAGTACGCAAGATCGGC 1740
QY 1741 ACCGCCACACCAAGAGACCTGAGACAGCTGACGAGGCGCTGCAAGAGATCGCCATGGAG 1800
DB 1741 ACCGCCACACCAAGAGACCTGAGACAGCTGACGAGGCGCTGCAAGAGATCGCCATGGAG 1800
QY 1794 ACCGCCACACCAAGAGACCTGAGACAGCTGACGAGGCGCTGCAAGAGATCGCCATGGAG 1794
QY 1801 AGATCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGAGACCTGG 1860
DB 1801 AGATCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGAGACCTGG 1860
QY 1861 GAGACCTGCTGGAGCCGACTACTGCGAGCCGACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
DB 1861 GAGACCTGCTGGAGCCGACTACTGCGAGCCGACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
QY 1921 ACCGCCCGCTGCTGAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCGGCGCCGAG 1980
DB 1921 ACCGCCCGCTGCTGAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCGGCGCCGAG 1980
QY 1981 ACCTTCTACCTGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCGGCTAGCTG 2040
DB 1981 ACCTTCTACCTGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCGGCTAGCTG 2040
QY 2041 ACCGACCGGCGCGGAGAGATCGTGAGCTGACCGAGACCCAGACCAAGAGAGAGAGAG 2100
DB 2041 ACCGACCGGCGCGGAGAGATCGTGAGCTGACCGAGACCCAGACCAAGAGAGAGAGAG 2100
QY 2101 CTGAGGCCATCCAGCTGGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
DB 2101 CTGAGGCCATCCAGCTGGGCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
QY 2161 AGCAGTACGCGCTGGGCTGAGTCCAGGCCAGCCCGAGAGAGAGAGAGAGAGAGAGAGAG 2220
DB 2161 AGCAGTACGCGCTGGGCTGAGTCCAGGCCAGCCCGAGAGAGAGAGAGAGAGAGAGAGAG 2220
QY 2221 AACAGATCATCGAGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
DB 2221 AACAGATCATCGAGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2280
QY 2281 CACAGGCGATCGGCGGACGAGCAGATCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
DB 2281 CACAGGCGATCGGCGGACGAGCAGATCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
QY 2341 GTGCTGTCTTGGAGCGCATCGATGGCGGCTGATGATCTACCAAGTACATGAGAGAGAGAG 2400
DB 2341 GTGCTGTCTTGGAGCGCATCGATGGCGGCTGATGATCTACCAAGTACATGAGAGAGAGAGAG 2400
QY 2401 TACGTGGGCGAGCGCGGCTTAGGATCGATTTAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
DB 2401 TACGTGGGCGAGCGCGGCTTAGGATCGATTTAAAGCTTCCCGGGGCTAGCACCGGTGAA 2460
QY 2461 TTC 2463
DB 2461 TTC 2457

RESULT 4
AX427930
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AX427930 9166 bp DNA linear PAT 20-JUN-2002
Sequence 168 from Patent WO0232943.

AX427930
AX427930.1 GI:21538017
synthetic construct.

ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 168 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1..9166
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2227 a 2662 c 2513 g 1764 t
ORIGIN

Query Match 82.3%; Score 2027.6; DB 6; Length 9166;
Best Local Similarity 91.7%; Pred. No. 5.5e-204;
Matches 2194; Conservative 0; Mismatches 179; Indels 19; Gaps 4;

QY 14 TGGCCGAGGCGATGAGCCAGGCGCACAGCCCAACATCTCTGATGCGAGCGCAGCAACTTCA 73
DB 2961 TGGCCGAGGCGATGAGCAAGGTGAACAACCAACCAACATCATGATGCGAGCGCAGCACTGCA 3020

QY 74 AGGCCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGCGCCACATCGCCCGCA 133
DB 3021 AGGCCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGGAGGCGCCACATCGCCCGCA 3080

QY 134 ACTGCGCGCGCCCGCAAGAGGCTGCTGGAAGTCCGCAAGGAGGCGCCACAGATGA 193
DB 3081 ACTGCGCGCGCCCGCAAGAGGCTGCTGGAAGTCCGCAAGGAGGCGCCACAGATGA 3140

QY 194 AGGACTGACCGAGCGACAGGCTAA-TTTTTTAGGGAAGATCTGGCTTCCCAAGGGA 3199
DB 3141 AGGACTGACCGAGCGACAGGCTAA-TTTTTTAGGGAAGATCTGGCTTCCCAAGGGA 3199

QY 254 AGGCCCGCGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCGCAGCGGAGCTGC 313
DB 3200 AGGCCCGCGAGTTCCCGAGCGAGCAGAACCGCGCCCAACAGCCCGCAGCGGAGCTGC 3259

QY 314 AGGTGCGCGG- ----CGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGCA --- 364
DB 3260 AGGTGCGCGGAGAGACAACTCCCTCTCAGAAGCAGGAGCGCGATAGACAAAGGAAGT 3319

QY 365 ----CCCTGAATTTCCCGAGATCACCTCTGCGAGCGCCCGCTGGTGAAGATCAAGTGG 421
DB 3320 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGATAG 3379

QY 422 GCGGCCAGATCAAGGAGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCGG 481
DB 3380 GGGGCCAGCTGAAGGAGCGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCG 3439

QY 482 TGAGCCTGCGCGCGCAAGTGAAGCCCAAGATGATCGGCGCGCATCGGCGGCTTCATCAAG 541
DB 3440 TGAACCTGCGCGCGCTGGAAGCCCAAGATGATCGGCGCGCATCGGCGGCTTCATCAAG 3499

QY 542 TGGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCGCGCTGC 601
DB 3500 TGGGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCGCGCTGC 3559

QY 602 TGATCGGCCCGCCCGCTGGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGTGCA 661
DB 3560 TGTGGGCGCCCGCGCTGGAACATCATCGGCGCGCAACATGCTGACCCAGATCGCGTGCA 3619

QY 662 CCCTGAATTTCCCGATCAGCGCCCATCGAGACCGTGCCTGGAAGCTGAAGCGCGCATGG 721
DB 3620 CCCTGAATTTCCCGATCAGCGCCCATCGAGACCGTGCCTGGAAGCTGAAGCGCGCATGG 3679

QY 722 ACGGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAGATCAAGGCGCTGACCGCCA 781
DB 3680 ACGGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAGATCAAGGCGCTGCTGGAGA 3739

QY 782 TCTGCGGAGGATGGAGAAGGAGGCGCAAGATCAACAAGATCGGCGCGCGCGCGCGCGCG 841

DB 3740 TCTGCACCGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCGCGAGAACCCTTACA 3799

QY 842 ACACCCCGCTGTTCCGCCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTGGACT 901
DB 3800 ACACCCCGCTGTTCCGCCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTGGACT 3859

QY 902 TCCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGTCCAGTGGGCAATCCCCAC 961
DB 3860 TCCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGTCCAGTGGGCAATCCCCAC 3919

QY 962 CCGCCGCGCTCAAGAAGAAAGAGCGTGCCTGGAGTGGGCGACGCTTACTTCA 1021
DB 3920 CCGCCGCGCTGAAGCAAGAAGAGCGTGCCTGGAGTGGGCGACGCTTACTTCA 3979

QY 1022 GCGTGCCTGCGGAGGAGGACTTCCGCAAGTACACCGGCTTCAACCATCCCCAGCATACA 1081
DB 3980 GCGTGCCTGCGGAGGAGGACTTCCGCAAGTACACCGGCTTCAACCATCCCCAGCATACA 4039

QY 1082 ACGAGACCCCGGATCCGCTACCAAGTACACGCTGCCCGAGGCTGGAAGGCGAGCC 1141
DB 4040 ACGAGACCCCGGATCCGCTACCAAGTACACGCTGCCCGAGGCTGGAAGGCGAGCC 4099

QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACCCG 1201
DB 4100 CCGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAACCCG 4159

QY 1202 AGATCGTGATCTACCA- ----GGCCCCCTGTACGTGGGCGAGGCTGGAGATCGGCC 1255
DB 4160 ACATCGTGATCTACCAAGTACATGACCAACCTGTACGTGGGCGAGGCTGGAGATCGGCC 4219

QY 1256 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAACCTGCTGCGTGGGCTTCAACACCC 1315
DB 4220 AGCACCGCAAGATCGAGGAGCTGCGCAAGCAACCTGCTGCGTGGGCTTCAACACCC 4279

QY 1316 CCGACAAGAGCACCAGAAGAGCCCGCTTCTCTGTGGATGGGCTAGAGCTGCACCCCG 1375
DB 4280 CCGACAAGAGCACCAGAAGAGCCCGCTTCTCTGTGGATGGGCTAGAGCTGCACCCCG 4339

QY 1376 ACAAGTGGACCGTGCAGCCCATCGAGTGCCTGCGCAAGAGGAGGCTGACCGTGAACACA 1435
DB 4340 ACAAGTGGACCGTGCAGCCCATCGTGTGCTGCGCAAGAGGAGGCTGACCGTGAACACA 4399

QY 1436 TCAGAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 1495
DB 4400 TCAGAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCGCGCATCAAGTGC 4459

QY 1496 GCAGCTGTCAAGCTGCTGCGGCGGCCAAGGCCCTTGACCGCATCGTGCCTGTGACCG 1555
DB 4460 GCAGCTGTCAAGCTGCTGCGGCGGCCAAGGCCCTTGACCGCATCGTGCCTGTGACCG 4519

QY 1556 AGGAGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCTCTGCGGAGCGCGTGCACGCG 1615
DB 4520 AGGAGCGGAGCTGGAGCTGGCGAGAACCGCGAGATCTCTGGAAGGAGCGCGTGCACGCG 4579

QY 1616 TGTACTACGACCCAGCAAGGAGCTGCTGGCGCGAGATCCAGAGCAGGCGCATCAAGCAGT 1675
DB 4580 TGTACTACGACCCAGCAAGGAGCTGATCCCGAGATCCAGAGCAGGCGCATCAAGCAGT 4639

QY 1676 GGACCTACAGATCTTACAGAGGCGCTTCAAGAACCTTGAAGCCGCGCAAGTACGCCAAGA 1735
DB 4640 GGACCTACAGATCTTACAGAGGCGCTTCAAGAACCTTGAAGCCGCGCAAGTACGCCGCA 4699

QY 1736 TGGCGACCGCCACACAGCAGTGAAGCAGCTGACCGAGGCGCTGCAGAAATGCCCA 1795
DB 4700 TGAAGGCGCGCCACACAGCAGTGAAGCAGCTGACCGAGGCGCTGCAGAAATGCCCA 4759

QY 1796 TGGAGGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCGCCCTGCCCATCCAGAAGAGA 1855
DB 4760 CCGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCAGCTGCCCATCCAGAAGAGA 4819

QY 1856 CPTGGAGACCTGGTGGAGCGACTACTGGCAGGCCCACTTGGATCCCGGAGTGGAGTTGC 1915

```
Db 4820 CTTGGAGGCGCTGGTGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTGC 4879
QY 1916 TGAACACCCCGCTGGTGAAGCTGTGTACCACTGGAGAGGAGGCCCATCATCGGCG 1975
Db 4880 TGAACACCCCGCTGGTGAAGCTGTGTACCACTGGAGAGGAGGCCCATCATCGGCG 4939
QY 1976 CCGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGACCAAGATCGGCAAGCGCGCT 2035
Db 4940 CCGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGACCAAGCTGGGCAAGCGCGCT 4999
QY 2036 ACGTAGCCAGCCGGCGCGCAGAAGATCTGTAGCCTGTGACCGAGACCAACCAAGAGA 2095
Db 5000 ACGTAGCCAGCCGGCGCGCAGAAGTGTGTGCTTACCGGACACCAACCAAGAGA 5059
QY 2096 CCGAGCTGCAGGCGCATCCAGCTGSCCTGCAGGACAGCGCAGGAGTGAACATCGTA 2155
Db 5060 CCGAGCTGCAGGCGCATCCAGCTGSCCTGCAGGACAGCGCCTGGAGGTGAACATCGTA 5119
QY 2156 CCGACAGCCAGTAGCGCCTGGGCATCATCCAGGCCCAAGCCGACAGAGGAGAGCGAGC 2215
Db 5120 CCGACAGCCAGTAGCGCCTGGGCATCATCCAGGCCCAAGCCGACAGAGGAGAGCGAGC 5179
QY 2216 TGGTGAACAGATCATCGAGCAGCTGTATCAAGAGGAGAGGTGTACTGAGCTGGGTGC 2275
Db 5180 TGGTGAACAGATCATCGAGCAGCTGTATCAAGAGGAGAGGTGTACTGAGCTGGGTGC 5239
QY 2276 CCGCCACAGAGGCGATCGCGGCAACAGCAGAGTGCAGACAGCTGGTGCAGCAAGGCGATCC 2335
Db 5240 CCGCCACAGAGGCGATCGCGGCAACAGCAGAGTGCAGACAGCTGGTGCAGCGCGGCGATCC 5299
QY 2336 GCAAGGTGCTTCTCTGGAGCGGCATGATGGCGCATCGTGATCTACCACTA 2387
Db 5300 GCAAGGTGCTTCTCTGGAGCGGCATGACACAGGCCCGCAGGAGGACGACAGAA 5351

RESULT 5
AX427936
LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 174 from Patent WO0232943.
ACCESSION AX427936
VERSION AX427936.1 GI:21538023
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1. Huang, Y. and Nabel, G. J.
   Modifications of hiv env, gag, and pol enhance immunogenicity for
   genetic immunization
   Patent: WO 0232943-A 174 25-APR-2002;
   GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source
1. .9788
   /organism="synthetic construct"
   /db_xref="taxon:32630"
   /notes="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2377 a 2817 c 2695 g 1899 t
ORIGIN
Query Match 82.0%; Score 2019.2; DB 6; Length 9788;
Best Local Similarity 91.6%; Pred. No. 4.1e-203;
Matches 2190; Conservative 0; Mismatches 178; Indels 24; Gaps 4;

QY 14 TGGCGGAGCCATGACCGCAGCCAGCCAGCCCAACATCTGTATCGAGCGGCAACTTCA 73
Db 2961 TGGCGGAGCCATGAGCAAGGTGACCAACACCAACATCATGTATGACGCGCAGCACTGCA 3020
QY 74 AGGCGCCCAAGCGCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCACATCGCCCGCA 133
Db 3021 AGGCGCCCAAGCGCACCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCACCTGCCCGCA 3080
QY 134 ACTGCCCGCGCCCGCCCAAGAGGCGCTGCTGGAAGTGCAGGAGGCGCCACAGATGA 193
Db 4155 ACATCGTATCTACCAAGTACATGGACCACTGTACGTGGGAGCGACCTGGAGATCGGCC 4214
```

```
Db 3081 ACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGCAGGAGGCGCACAGGTGA 3140
QY 194 AGGACTGCACCGAGCGCCAGGCCCAACTTCTTCCGCGAGGAGCTTGGCCCTTCCCCAGGCA 253
Db 3141 AGGACTGCACCGAGCGCCAGAGGCTAA-----TAGGGAAGATCTTGGCCTTCCCAAGAGGA 3194
QY 254 AGGCGCCCGAGTTTCCCGAGCGCAGAGCAAGCGCGCCCAACAGCCCAACAGCGCGAGGTGC 313
Db 3195 AGGCGAGGGAATTTCTTCAGAGCAGACAGAGCCACAGCGCCCAACAGAGAGAGGTTC 3254
QY 314 AGGTGCGCGG-----CGACAACCCCGCAGCAGGCGCGCGCGCGCGCGAGCGCGAGGCA--- 364
Db 3255 AGGTTTGGGAAGAGACAACAATCTCTCAGAAAGAGGAGCGGATAGACAAGAACTG 3314
QY 365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTCGTGAGCATCAAGGTG 421
Db 3315 TATCTTTAGCTTCCCTTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAAGATAG 3374
QY 422 GCGGCCAGATCAAGGAGGCGCTGTGTGACACCGCGCGCGAGCACACCGTGTGGAGGAGA 481
Db 3375 GGGSCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGAGCACACCGTGTGGAGGAGA 3434
QY 482 TGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGCGATCGCGGCTTCAATCAGG 541
Db 3435 TGAACCTGCCCGCGCTGGAAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCAATCAGG 3494
QY 542 TGGCCAGTAGTACGACAGATCCTGTATCGAGATCTCGGCAAGAGCGCATCGCACCGTGC 601
Db 3495 TGGCCAGTAGTACGACAGATCCTGTATCGAGATCTCGGCGCGCAAGGCCATCGCACCGTGC 3554
QY 602 TGATCGCGCCCGCACCGCTGAACATCATCGCGCGCAACATCTGACCCAGCTGGGCTGCA 661
Db 3555 TGGTGGCGCCCGCACCGCTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCTGCA 3614
QY 662 CCCTGAACTTCCCGCATCAGCCCCATCGAGACCGTGCCTGCGTGAAGCTGAAGCCCGCATGG 721
Db 3615 CCCTGAACTTCCCGCATCAGCCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGCATGG 3674
QY 722 ACGGCCCCAAGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCGGCA 781
Db 3675 ACGGCCCCAAGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGCTGGGAGA 3734
QY 782 TCTCGAGGAGATGGAGAAGGAGGCGAAGATCAACCAAGATCGGCCCGCGAGAACCCCTACA 841
Db 3735 TCTGCACCGAGATGGAGAAGGAGGCGAAGATCAACCAAGATCGGCCCGCGAGAACCCCTACA 3794
QY 842 ACACCCCGCTTTCGCCATCAAGAGAGGAGACAGCAACCAAGTGGCGCAAGCTGTGGACT 901
Db 3795 ACACCCCGCTTTCGCCATCAAGAGAGGAGGAGACAGCAACCAAGTGGCGCAAGCTGTGGACT 3854
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCGATCCCCACC 961
Db 3855 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCGATCCCCACC 3914
QY 962 CCGCGCGCTTGAAGAAGAGAGCGCTGACCGCTGCTGGAGGTGGCGACCGCTACTTCA 1021
Db 3915 CCGCGCGCTTGAAGAAGAGAGAGCGCTGACCGCTGCTGGAGGTGGCGACCGCTACTTCA 3974
QY 1022 CGGTGCCCTCGGACGAGGACTTCCGCAAGTACACCGCTTCAACATCCCCAGCATCAACA 1081
Db 3975 CGGTGCCCTCGGACGAGGACTTCCGCAAGTACACCGCTTCAACATCCCCAGCATCAACA 4034
QY 1082 ACAGAGCCCCGGGATCGGCTACCAAGTACACCGCTTCAACATCCCCAGCATCAACA 1141
Db 4035 ACAGAGCCCCGGGATCGGCTACCAAGTACACCGCTTCAACATCCCCAGCATCAACA 4094
QY 1142 CCAGCATCTTCCAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCGCCGCAACCCCG 1201
Db 4095 CCGCATCTTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCGCGCAACCCCG 4154
QY 1202 AGATCGTATCTACCA-----GGCCCCCTTGTAGCTGGGAGCGACCTGGAGATCGGCC 1255
Db 4155 ACATCGTATCTACCAAGTACATGGACCACTGTACGTGGGAGCGACCTGGAGATCGGCC 4214
```


Best Local Similarity 91.3%; Pred. No. 2.8e-201; Matches 2186; Conservative 0; Mismatches 187; Indels 22; Gaps 5;									
Qy	14	TGGCGGAGCCATGAGCCAGG- --CCACAGCGCAACATCTGATGCGAGCGAGCAACT	70						
Db	2967	TGGCGGAGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGGCAACT	3026						
Qy	71	TCAAGGGCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGGCCACATCGCCC	130						
Db	3027	TCCGCAACAGCGCAAGATCGTGAAGTGTCAACTGCGGCAAGGAGGCCACACCGCCC	3086						
Qy	131	GCAACTGCGCGCCCGCCGCGCAAGAGGGCTGCTGGAAGTGGCGGAAGAGGGCCACACAGA	190						
Db	3087	GCAACTGCGCGCCCGCCGCGCAAGAGGGCTGCTGGAAGTGGCGGAAGGAGGGCCACACAGA	3146						
Qy	191	TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCCAGG	250						
Db	3147	TGAAGGACTGCACCGAGCGACAGGCTRA- TTTTATGGAAGATCTGGCCCTTCCCAAG	3205						
Qy	251	GCAAGGCCGCGGAGTTCCCCAGCGAGCAGAGAACCGCGCCACACAGCCCAAGCGCGAGC	310						
Db	3206	GGAAGGCCAGGGAATTTCTTTCAGAGCAGACACAGAGCCACAGCCACAGAGAGAGC	3265						
Qy	311	TGCAGGTGCGCG- ----CGACAAACCCCGCAGCGAGGCGCGGCGCGAGCGCCAGGGCA	364						
Db	3266	TTCAAGGTTTGGGGAAGAGACAACAACACTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGAA	3325						
Qy	365	-----CCCTGAACTTCCCCAGATCACTCTGTGGCAGCGCCCTCTGTGAGCATCAAGG	418						
Db	3326	CTGTATCTTTCCTTCCCTCAGTACTCTTTGGCAGCAGACCCCTCGTCACAATAAGA	3395						
Qy	419	TGGCGGCGCAGATCAAGAGGCGCCTGTGTGACACCGCGCGCAGACACCGTGTGTGAGG	478						
Db	3386	TAGGGGCGCAGCTGAAGGAGGCGCCTTCTAGACACCGCGCGCAGACACCGTGTGTGAGG	3445						
Qy	479	AGATGAGCCTGCCCGCAAGTGNAGCCCAAGATGATCGGGGGATCGGGCGCTTATCA	538						
Db	3446	AGATGAACTTGCCCGCGCGTGGAAAGCCCAAGATGATCGGGCGCATCGGGCGCTTCA	3505						
Qy	539	AGGTGCGCCAGTAGCAGACAGATCTGTATCGAGATCTCGGCGCAAGAGCCATCGGCACCG	598						
Db	3506	AGGTGCGCCAGTAGCAGACAGATCTGTATCGAGATCTCGGCGCACAGGCCATCGGCACCG	3565						
Qy	599	TGCTGATCGGCGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT	658						
Db	3566	TGCTGTGGGCGCCACCCCGTGAACATCATCGGCGCGCAACCTGCTGACCCAGATCGGCT	3625						
Qy	659	GCACCTGAATTCCTCCATCAGCCCATCGAGCCGTGCGCGTGAAGCTGAAGCCCGGCA	718						
Db	3626	GCACCTGAATTCCTCCATCAGCCCATCGAGCCGTGCGCGTGAAGCTGAAGCCCGGCA	3685						
Qy	719	TGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG	778						
Db	3686	TGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGTTGG	3745						
Qy	779	CCATCTCGAGAGGATGGAGAAGAGGGCAAGATCAACCAAGATCGGCGCCCGAGAACCCCT	838						
Db	3746	AGATCTGCACCGAGATGGAGAAGAGGGCAAGATCAGCAAGATCGGCGCCCGAGAACCCCT	3805						
Qy	839	ACAACACCCCGTTCGCGCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTGG	898						
Db	3806	ACAACACCCCGTTCGCGCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTGG	3865						
Qy	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCC	958						
Db	3866	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCC	3925						
Qy	959	ACCCCGCGCCTGAGAAGAAGAAGAGCTGACCGTGTGTGGACGTGGGCGACCGCTACT	1018						
Db	3926	ACCCCGCGCCTGAGAAGAAGAAGAGCTGACCGTGTGTGGACGTGGGCGACCGCTACT	3985						
Qy	1019	TCAGCGTGCCTGGAGGAGCTTCGCAAGTACACCGCCTTCACCATCCCCAGCATCA	1078						
Db									
Db	3986	TCAGCGTGCCTGGCAAAAGGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA	4045						
Qy	1079	ACAAGAGACCCCGGATCCGCTACCACTACAAGTCTGCTGCCAGGCTGGAGGGCA	1138						
Db	4046	ACAAGAGACCCCGGATCCGCTACCACTACAAGTCTGCTGCCAGGCTGGAGGGCA	4105						
Qy	1139	GCCCCAGCATTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCCGCAACC	1198						
Db	4106	GCCCCAGCATTTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCGCAAGCAGAACC	4165						
Qy	1199	CCGAGATCGTATCTACCA- ----GGCCCCCTGTAGCTGGGAGCAGCTGGAGATCG	1252						
Db	4166	CCGACATCGTATCTACCACTACATGGACACCTGTAGCTGGCAGCAGCTGGAGATCG	4225						
Qy	1253	GCCAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA	1312						
Db	4226	GCCAGCAGCGCACCAAGATCGAGGAGCTGCGCCAGCAGCTGCTGCTGGGGCTTCACCA	4285						
Qy	1313	CCCCCGACAAGAACCACAGAAGGAGCCCTTCTTCTGTGTGGATGGGTACGAGTGCACC	1372						
Db	4286	CCCCCGACAAGAAGCACCAAGAAGGAGCCCTTCTTCTGTGTGGATGGGTACGAGTGCACC	4345						
Qy	1373	CCGACAAGTGAGCCTGCAGCCCATCGAGCTGCGCGAGAGAGAGAGCTGGACCGTGAACG	1432						
Db	4346	CCGACAAGTGAGCCTGCAGCCCATCGCTGCTGCGGCAAGCTGAGCTGGCCGGCATCAAG	4405						
Qy	1433	ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCAAGCAGCATCTACCCCGGCATCAAG	1492						
Db	4406	ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCAAGCAGCATCTACCCCGGCATCAAG	4465						
Qy	1493	TGCGCCAGCTGTGAAGCTGCTGCGCGCGCCAAAGGCGCTTACCGAGCATCTGCGCCCTGA	1552						
Db	4466	TGCGCCAGCTGTGAAGCTGCTGCGCGGCAAGGCGCTTACCGAGCATCTGCGCCCTGA	4525						
Qy	1553	CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGGCGGAGCCCGTGCACG	1612						
Db	4526	CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCCTGGCGGAGCCCGTGCACG	4585						
Qy	1613	CGGTGTACTAGACCCAGCAAGGAGCTGGTGGCGGAGATCCAGAACGAGGCGGCAGACC	1672						
Db	4586	CGGTGTACTAGACCCAGCAAGGAGCTGGTGGCGGAGATCCAGAACGAGGCGGCAGGCC	4645						
Qy	1673	AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA	1732						
Db	4646	AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCC	4705						
Qy	1733	AGATGCCACCGCCACCAACAGCAGTGAAGAGCTGACCGAGGCGCTGCAGAGATCG	1792						
Db	4706	GCATGAAGGCGGCCACCAACAGCAGTGAAGCAGCTGACCGAGGCGCTGCAGAGATCG	4765						
Qy	1793	CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTTCGCCCTGCCATCCAGAGG	1852						
Db	4766	CCACCGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTTCAGCTGCCATCCAGAGG	4825						
Qy	1853	AGACCTGGGAGACCTGGTGGACCGCACTACTGGCAGGCGACCTGGATCCCGAGTGGAGT	1912						
Db	4826	AGACCTGGGAGCCTGGTGGACCGAGTACTGGCAGGCGCACTGGATCCCGAGTGGAGT	4885						
Qy	1913	TCGTGAACACCCCGCCCTGGTGAAGCTGTGTATACAGCTGGAGAGAGAGCCCATCATCG	1972						
Db	4886	TCGTGAACACCCCGCCCTGGTGAAGCTGTGTATACAGCTGGAGAGAGAGCCCATCATCG	4945						
Qy	1973	GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCAGAGACCAAGATGGCAAGGCCG	2032						
Db	4946	GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCAGAGACCAAGATGGCAAGGCCG	5005						
Qy	2033	GCTACCTGACCGACCGGGCGCGCAAGATCTGTGAGCCTGACCGAGACCAACCAACAGA	2092						
Db	5006	GCTACCTGACCGACCGGGCGCGCAAGATCTGTGAGCCTGACCGAGACCAACCAACAGA	5065						
Qy	2093	AGACCGAGCTGCAGGCGCATCCAGCTGCCCTTCAGGAGACAGGCGGAGGTGAACATCG	2152						
Db	5066	AGACCGAGCTGCAGGCGCATCCAGCTGCCCTTCAGGAGACAGGCGGCTGGAGGTGAACATCG	5125						

QY 2153 TGACGACAGCCAGTAGCCCTGGGATCATCCAGGCCAGCCGACAAAGAGGAGAGCG 2212
|||||
Db 5126 TGACGACAGCCAGTAGCCCTGGGATCATCCAGGCCAGCCGACAAAGAGGAGAGCG 5185
|||||
QY 2213 AGCTGGTGAACACAGATCATCGACGAGTGTATCAAGAGGAGAGAGTGTACCTGAGCTGGG 2272
|||||
Db 5186 AGCTGGTGAACACAGATCATCGACGAGTGTATCAAGAGGAGAGAGTGTACCTGAGCTGGG 5245
|||||
QY 2273 TGCCGCCACACAGGATCGGCGGCGCAACGACGAGATCGACAGCTGGTGAAGGCGCA 2332
|||||
Db 5246 TGCCGCCACACAGGATCGGCGGCGCAACGACGAGTGGAGCGCTGGTGAAGGCGCGCA 5305
|||||
QY 2333 TCCGCAAGGTGCTGTTCTTGGAGCGGATCGATGGCGGATCGTGTATACAGTA 2387
|||||
Db 5306 TCCGCAAGGTGCTGTTCTTGGAGCGGATCGACAAAGGCCAGGAGGACAGCAAA 5360
|||||
RESULT 8
AX427925
LOCUS AX427925 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 163 from Patent WO0232943.
ACCESSION AX427925
VERSION AX427925.1 GI:21538012
SOURCE synthetic construct.
ORGANISM synthetic construct.
artificial sequences.
1
REFERENCE
AUTHORS Huang, Y. and Nabel, G. J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 163 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1..9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2671 c 2524 g 1767 t
ORIGIN
Query Match 81.2%; Score 2000.2; DB 6; Length 9194;
Best Local Similarity 91.2%; Pred. No. 4.1e-201;
Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;
QY 14 TGGCGAGGCCATGAGCCAGG---CCACAGCGGCACATCTGTATCGAGCGGCAACT 70
Db 2967 TGGCGAGGCCATGAGCCAGGTCACCAACAGCGCCACCATCATGATCGAGCGGCAACT 3026
QY 71 TCAAGGSCCCAAAGCGCATCATCAAGTCTCAACTCGGCGAAGGAGGCGCACATCGCCC 130
Db 3027 TCAGCAACAGCGCAAGATCGTGAAGTCTCAACTCGGCGAAGGAGGCGCACACCGCCC 3086
QY 131 GCAACTCGCGCGCCCGCCGCAAGAAGGGCTGCTGGAAGTGGCGAAGGAGGCGCCACAGA 190
Db 3087 GCAACTCGCGCGCCCGCCGCAAGAAGGGCTGCTGGAAGTGGCGAAGGAGGCGCCACAGA 3146
QY 191 TGAAGGACTGACGAGCGCGCCAGCCAACTCTTCCCGAGGAGGAGGAGGAGGAGGAGG 250
Db 3147 TGAAGGACTGACGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3205
QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 310
Db 3206 GGAAGGCGAGGAGATTTCTTCAGAGCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGG 3255
QY 311 TGCAGGTCGCGGG-----CGAACACCCCGCAGCGAGGCGGCGGAGCGCGGAGGCGCA 364
Db 3266 TTCAGGTTTGGGGAAGAGAACAACTCCCTCTCAGAGGAGGAGGAGGAGGAGGAGGAGG 3325
QY 365 -----CCCTGAACCTTCCCGCAGATCACTCTGTGGAGCGCGCCCTGTGTGAGCATCAAG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGTGTGAGCATCAAG 4465

QY 419 TGGCGGCGCAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 478
|||||
Db 3386 TAGGGGCGCAGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3445
|||||
QY 479 AGATGAGCTTGGCGGCGCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 538
|||||
Db 3446 AGATGAGCTTGGCGGCGCGCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3505
|||||
QY 539 AGTGCGCGCAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 598
|||||
Db 3506 AGTGCGCGCAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3565
|||||
QY 599 TGTGATGCGGCGCCACCGCTGAAACATCATCGGCGCGCAACATGCTGACCCAGCTGGCT 658
|||||
Db 3566 TGTGATGCGGCGCCACCGCTGAAACATCATCGGCGCGCAACATGCTGACCCAGCTGGCT 3625
|||||
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 718
|||||
Db 3626 GCACCTGAACTTCCCATCAGCCCATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3685
|||||
QY 719 TGGAGGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGGAGGAGGAGGAGGAGG 778
|||||
Db 3686 TGGAGGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGGAGGAGGAGGAGGAGG 3745
|||||
QY 779 CCATCTCGAGGAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838
|||||
Db 3746 AGATCTGACCGAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3805
|||||
QY 839 ACAACACCCCGTGTTCGCCATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898
|||||
Db 3806 ACAACACCCCGTGTTCGCCATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3865
|||||
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 958
|||||
Db 3866 ACTTCCGCGAGCTGAACAAGCGCACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3925
|||||
QY 959 ACCCGCGCGGCTGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1018
|||||
Db 3926 ACCCGCGCGGCTGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3985
|||||
QY 1019 TCAGCTGCGCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1078
|||||
Db 3986 TCAGCTGCGCCCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4045
|||||
QY 1079 ACACGAGAGCCCGCGCATCCGCTACCAAGTACAAAGTGTCTGCGGAGGAGGAGGAGG 1138
|||||
Db 4046 ACACGAGAGCCCGCGCATCCGCTACCAAGTACAAAGTGTCTGCGGAGGAGGAGGAGG 4105
|||||
QY 1139 GCGCGCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGGAGGAGGAGGAGGAGG 1198
|||||
Db 4106 GCGCGCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGGAGGAGGAGGAGGAGG 4165
|||||
QY 1199 CCGAGATCGTGAATACCA-----GGCGCCCTGTAGTGGGAGGAGGAGGAGGAGGAGG 1252
|||||
Db 4166 CCGAGATCGTGAATACCA-----GGCGCCCTGTAGTGGGAGGAGGAGGAGGAGGAGGAGG 4225
|||||
QY 1253 GCGAGAGCCCGCGCAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1312
|||||
Db 4226 GCGAGAGCCCGCGCAAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4285
|||||
QY 1313 CCGCGCAGCAAGAGCAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1372
|||||
Db 4286 CCGCGCAGCAAGAGCAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4345
|||||
QY 1373 CCGACAGTGAAGCTGACGCGCATCGAGCTGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGG 1432
|||||
Db 4346 CCGACAGTGAAGCTGACGCGCATCGAGCTGCGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4405
|||||
QY 1433 ACATCCAGAGGAGTGGTGGGAGGAGTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1492
|||||
Db 4406 ACATCCAGAGGAGTGGTGGGAGGAGTGAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4465

QY 1493 TGGCCAGCTGTGCAAGCTGTGTGGCGCGCCCAAGGCCCTGACCGACATCGTGGCCCTGA 1552
Db TGGCCAGCTGTGCAAGCTGTGTGGCGCGCCCAAGGCCCTGACCGAGGTGTGGCCCTGA 4525
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGGCGGACCGCGTGCACG 1612
Db CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAAGGAGCGCGTGCACG 4585
QY 1613 GCGTGTACTAGGACCCAGCAAGACCTGTGTGGCGGAGATCCAGAACGAGCGCCACGACC 1672
Db GCGTGTACTAGGACCCAGCAAGACCTGTGTGGCGGAGATCCAGAACGAGCGCGAGGCCC 4645
QY 1673 AGTGGACCTACCAAGCTTACCAAGAGCCCTTCAAGAACCCTGAAGACCGCAAGTACGCCA 1732
Db AGTGGACCTACCAAGCTTACCAAGAGCCCTTCAAGAACCCTGAAGACCGCGCAAGTACGCC 4705
QY 1733 AGATGCCACCGCCACCAACAGACGCTGAAGACGCTGACCGAGGCGCTGCAGAAGATCG 1792
Db AGATGCCACCGCCACCAACAGACGCTGAAGACGCTGACCGAGGCGCTGCAGAAGATCG 4765
QY 1793 CCATGGAGAGCATCGTGTGTGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAGG 1852
Db CCATGGAGAGCATCGTGTGTGGGCAAGACCCCAAGTTCCAGCTGCCCATCCAGAGG 4825
QY 1853 AGACCTGGGAGACCTGTGTGACCGACTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 1912
Db AGACCTGGGAGGCTGTGTGACCGAGTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT 4885
QY 1913 TCGTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGGAGGCGCATCATCG 1972
Db TCGTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGGAGGCGCATCATCG 4945
QY 1973 GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAAGCCGAGACCAAGATCGCAAGGCCG 2032
Db GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAAGCCGAGACCAAGATCGCAAGGCCG 5005
QY 2033 GCTACGTGACCGAGCGCGCGCGCGAGAGATCGTGTGACCTGACCGAGACCAACACAGA 2092
Db GCTACGTGACCGAGCGCGCGCGCGAGAGTGTGTGACCTGACCGAGACCAACACAGA 5065
QY 2093 AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACACGCGCGAGGTGAACATCG 2152
Db AGACCGAGCTGACGAGCATCCAGCTGGCCCTGCAGGACACGCGCGAGGTGAACATCG 5125
QY 2153 TGACCGACAGCCAGTACGCCCTGGGATCATCCAGGCGCCAGCCGCGAGACGCGAGGTGAACATCG 2212
Db TGACCGACAGCCAGTACGCCCTGGGATCATCCAGGCGCCAGCCGCGAGACGCGAGGTGAACATCG 5185
QY 2213 AGCTGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2272
Db AGCTGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 5245
QY 2273 TGCCCGCCCAAGGCGATCGCGCGCAACGAGCAGATCGACAGCTGGTGAAGAGGGCA 2332
Db TGCCCGCCCAAGGCGATCGCGCGCAACGAGCAGGTGGACGCGCTGGTGAAGAGGGCA 5305
QY 2333 TCCGCAAGGTGCTTCTTCCGAGGCGATCGATGGCGGCGATCGTGTACCACTA 2387
Db TCCGCAAGGTGCTTCTTCCGAGGCGATCGACAGGCGCCAGGAGGAGCAGAGAA 5360

RESULT 9
AX427927
LOCUS AX427927 12411 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 165 from Patent W00232943.
ACCESSION AX427927
VERSION AX427927.1 GI:21538014
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
JOURNAL genetic immunization
FEATURES Patent: WO 0232943-A 165 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
Location/Qualifiers
source 1. .12411
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2938 a 3733 c 3476 g 2264 t
ORIGIN
Query Match 81.2%; Score 2000.2; DB 6; Length 12411;
Best Local Similarity 91.2%; Pred. No. 3.8e-201;
Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTCTGATGACGCGCAGCAACT 70
Db TGGCCGAGGCGCATGAGCCAGGTTGACCAACAGCGCCACCATCATCATGACGCGCGCAACT 3026
QY 71 TCAAGGCGCCCAAGCGGATCATCAAGTGTCTCAACTGCGGCAAGAGGGCCACATCGGCC 130
Db TCCGCAACACGCGCAAGATCGTGAAGTGTCTCAACTGCGGCAAGAGGGCCACACCGGCC 3086
QY 131 GCAACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 190
Db GCAACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 3146
QY 191 TGAAGGACTCACCGAGCGCGCCAGCCAACTTCTTCCCGAGGACCTTGGCCCTTCCCCCAGG 250
Db TGAAGGACTCACCGAGCGAGGCTAA-TTTTATAGGAAGATCTGGCCTTCCCCACAAG 3205
QY 251 GCAAGGCGCCCGAGTTCGCCAGCGAGCAGAGACCGCGCCACACGCCCCACAGCGCGCAGC 310
Db GGAAGGCGCCAGGAAATTTCTTCAGAGCAGACAGACAGCCACAGCCCCACAGAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCAGCGCA 364
Db TTCAAGTTTGGGGAAGAGACAACTCCCTCTCAGAGCAGGAGCGCGATAGACAGGAA 3325
QY 365 -----CCCTGAACCTTCCCGAGATCACTCTGTGGCAGCGCGCCCTGGTGAGCATCAAGG 418
Db CTGTATCTCTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAGA 3385
QY 419 TGGCGCGCGAGATCAAGAGGCGCTCTGTGACACCGCGCGCAGACACCGCTGTGAGG 478
Db TAGGGGCGAGCTGAGAGGAGCGCTCTTAGACACCGCGCGCGCAGACACCGCTGTGAGG 3445
QY 479 AGATGAGCCTTCCCGCGCAAGTGAAGCCCAAGATGATCGCGCGCGCATCGCGCGCTTCATCA 538
Db AGATGAACCTTCCCGCGCGCTGGAAGCCCAAGATGATCGCGCGCGCATCGCGCGCTTCATCA 3505
QY 539 AGTGCGCCAGTACGACCAAGTCTGATCAGATCTGCGGCAAGAGGCGCATCGGCGACG 598
Db AGTGCGCCAGTACGACCAAGTCTGATCAGATCTGCGGCAAGAGGCGCATCGGCGACG 3565
QY 599 TGTGTGATCGCGCCCGCGCTGACATCATCGCGCGCACACCTGCTGACCCAGCTGGCT 658
Db TGTGTGATCGCGCCCGCGCTGACATCATCGCGCGCACACCTGCTGACCCAGCTGGCT 3625
QY 659 GCACCTGAACCTTCCCGCATCAGCCCGCATCGAGACCGCTGCGCGTGAAGCTGAAGCGCGCA 718
Db GCACCTGAACCTTCCCGCATCAGCCCGCATCGAGACCGCTGCGCGTGAAGCTGAAGCGCGCA 3685
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGAGTGG 3745
QY 779 CCATCTCGAGGAGAGTGGAGAGGAGGAGATCACCAGATCGCGCGCGCGCGAGAACCCCT 838
Db AGATCTGACCGAGATGGAGAGGAGGAGATCAGAGATCGCGCGCGCGCGAGAACCCCT 3805
QY 839 ACAACACCCCGCTGTTCGCCATCAAGAGAGGACAGACCAAGTGGCGCGCAAGCTGGTG 898

Db 3806 ACAACACCCCGCTGTTCGCCATCAAGAAAGACAGCACCAAGTGGCGCAAGCTGGTG 3865
QY 899 ACTTTCGGGAGCTGAACAAAGCGACCCAGAGCTTCTGGGAGGTGAGCTGGGATCCCC 958
Db 3866 ACTTTCGGGAGCTGAACAAAGCGACCCAGAGCTTCTGGGAGGTGAGCTGGGATCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAAGAGAGCGGTGACCGCTGTGGAGCTGGCGAGCGCTACT 1018
Db 3926 ACCCGCGCGCTGAAGCAAGAGAGCGGTGACCGCTGTGGAGCTGGCGAGCGCTACT 3985
QY 1019 TCAGGCTGCCCCGTGACGAGGAGCTTCCGGAAGTACACCGCTTACCATCCCAAGATCA 1078
Db 3986 TCAGGCTGCCCCGTGACGAGGAGCTTCCGGAAGTACACCGCTTACCATCCCAAGATCA 4045
QY 1079 ACAAGAGACCCCGCGATCCCGCTACCACTACAAGCTGTGCCCGAGGCTGGAAGGCA 1138
Db 4046 ACAAGAGACCCCGCGATCCCGCTACCACTACAAGCTGTGCCCGAGGCTGGAAGGCA 4105
QY 1139 GCCCAGACNTTCCAGAGCAGATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACC 1198
Db 4106 GCCCGCCATCTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CCGAGATCGTGATCTACCA- ----GGCCCCCTGTACGTGGGAGCGACCTGGAGATCG 1252
Db 4166 CCGATCTGTATCTACCAAGTACATGGAGCGACCTGTACGTGGGAGCGACCTGGAGATCG 4225
QY 1253 GCCAGACCGCCCAAGATCGAGGAGCTGCGCAAGCACTGTGCGCTGGGCTTTCACCA 1312
Db 4226 GCCAGACCGCCCAAGATCGAGGAGCTGCGCAAGCACTGTGCGCTGGGCTTTCACCA 4285
QY 1313 CCCCCAGAACGACCAAGAGAGCGCCCTTCTGTGGATGGGCTACGAGCTGCACC 1372
Db 4286 CCCCCAGAACGACCAAGAGAGCGCCCTTCTGTGGATGGGCTACGAGCTGCACC 4345
QY 1373 CGGACAAGTGAACCTGAGCGCCATCGAGCTGCCGAGAGAGAGAGCTGGACCTGAAG 1432
Db 4346 CCGACAAGTGAACCTGAGCGCCATCGTGTGCCGAGAGAGAGAGCTGGACCTGAAG 4405
QY 1433 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGCGGACGACATCTACCGCGCATCAAG 1492
Db 4406 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGCGGACGACATCTACCGCGCATCAAG 4465
QY 1493 TCGCCAGCTGTGAAGCTGTGCGCGCGGCAAGCCCTGACGACATCTGTCGCCCTGA 1552
Db 4466 TCGCCAGCTGTGAAGCTGTGCGCGGCAAGCCCTGACGAGCTGTGTCGCCCTGA 4525
QY 1553 CCGAGGAGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCCCGTGCAG 1612
Db 4526 CCGAGGAGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGAGGAGCCCGTGCAG 4585
QY 1613 GCGTGTACTAGACCCCAAGCAAGCACTGTGGCGGAGATCCAGAACGAGGCGCCAGCAC 1672
Db 4586 GCGTGTACTAGACCCCAAGCAAGCACTGTGGCGGAGATCCAGAACGAGGCGCCAG 4645
QY 1673 AGTGGACCTACAGATCTACAGAGCGCTTCAAGACCTGAAGACCGGCGAGTACGCCA 1732
Db 4646 AGTGGACCTACAGATCTACAGAGCGCTTCAAGACCTGAAGACCGGCGAGTACGCC 4705
QY 1733 AGATGCGCACCCGACACCAAGCAAGCTGAAGAGCTGACCGAGCGCTGCGAGAGATCG 1792
Db 4706 GCATGAAGGCGCCCAACCAAGCAAGCTGAAGAGCTGACCGAGCGCTGCGAGAGATCG 4765
QY 1793 CCATGGAGAGATCGTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAG 1852
Db 4766 CCACGAGAGATCGTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAG 4825
QY 1853 AGACCTGGAGAGCTGTGACCAAGCTACTGGAGGCGCACTGTGATCCCCGAGTGGAGT 1912
Db 4826 AGACCTGGAGAGCTGTGACCAAGCTACTGGAGGCGCACTGTGATCCCCGAGTGGAGT 4885
QY 1913 TCGTGAACACCCCGCTGTGTAAGCTGGTACCAAGCTGGAGAGGAGGCGCCATCATCG 1972

Db 4886 TCGTGAACACCCCGCTGTGTAAGCTGGTACCAAGCTGGAGAGGAGGCGCCATCATCG 4945
QY 1973 GCCTCCGAGACCTTCTACGTGGAGCGGCCCAACCGCGAGACCAAGATCGGCAAGGCGC 2032
Db 4946 GCCTCCGAGACCTTCTACGTGGAGCGGCCCAACCGCGAGACCAAGTGGGCAAGGCGC 5005
QY 2033 GCTACGTGACCGAGCGGCGCGCAGAGATCTGAGCCCTGACCGAGACCAACAGCA 2092
Db 5006 GCTACGTGACCGAGCGGCGCGCAGAGGTGTGTCCTGACCGACCAACAGCAACAGCA 5065
QY 2093 AGACCGAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCG 2152
Db 5066 AGACCGAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGGCTGGAGTGAACATCG 5125
QY 2153 TGACCGACAGCAGTACGAGCCCTGGGCGATCATCCAGGCGCCAGCGAGGAGAGCG 2212
Db 5126 TGACCGACAGCAGTACGAGCCCTGGGCGATCATCCAGGCGCCAGCGAGGAGAGCG 5185
QY 2213 AGCTGTGTAACAGATCATCGAGCAGCTGATCAAGAGAGAGAGGTGTACCTGAGCTGG 2272
Db 5186 AGCTGTGTAACAGATCATCGAGCAGCTGATCAAGAGAGAGAGGTGTACCTGAGCTGG 5245
QY 2273 TGCCCGCCCAAGGCGATCGGCGCAAGCAGATCGAGCAGCTGGTGAAGAGGCA 2332
Db 5246 TGCCCGCCCAAGGCGATCGGCGCAAGCAGAGCTGGAGCGCTGGTGAAGAGGCA 5305
QY 2333 TCCGCAAGTGTCTTCTGGAGCGGATCGAGCGGCGATCGTGTATCTACCAAGTA 2387
Db 5306 TCCGCAAGTGTCTTCTGGAGCGGATCGAGCGGCGATCGAGCGGCGAGGAGGAGCA 5360

RESULT 10
AX427938
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 176 from Patent WO0232943.
ACCESSION AX427938
VERSION AX427938.1 GI:21538025
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G. J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 176 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
1..9785
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2360 a 2843 c 2731 g 1851 t
ORIGIN

Query Match 81.1%; Score 1996.8; DB 6; Length 9785;
Best Local Similarity 91.3%; Pred. No. 9.1e-201;
Matches 2183; Conservative 0; Mismatches 182; Indels 27; Gaps 5;

QY 14 TGGCGAGGCGCATGAGCCAGGCGCACCATCTGATGCGAGCGAGCAACTTCA 73
Db 2967 TGGCGAGGCGCATGAGCCAGGTGCGAGCACACCAATCATGATGAGCGGCACTTCA 3026
QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTCACTGCGGCAAGAGGCGCACATCGCCCCGA 133
Db 3027 AGGCGCCCAAGCG---CATCAAGTGTCTCACTGCGGCAAGAGGCGCACCTTGGCCGCA 3083
QY 134 ACTGCGCGCGCCCGCGCAAGAGGCTGCTGGAAGTGGCGAAGAGGAGGCGCACCATGGA 193
Db 3084 ACTGCGCGCGCTCGGCAAGAGGCTGCTGGAAGTGGCGAAGAGGAGGCGCACCATGGA 3143
QY 194 AGGACTGCGAGCGCGCGCAAGGCAACTTCTTCCGCGAGGAGCTGGCTTCCCGCGAGGCA 253

Db 3144 AGGACTGCACCGGACAGGCTAA-----TAGGGAAGATCTGGCCCTTCCCAACAGGGA 3197
Qy 254 AGGCCCGGAGTTCCTCCAGGAGCAGAAACCGCGCCAAACAGCCCAACACAGCCGCGAGTGC 313
Db 3198 AGGCCAGGGAATTTCTTCAGAGCAGACAGACAGCAACAGCCCAACAGAGAGAGCTTC 3257
Qy 314 AGTGGCCGG-----CGACAACCCCGCAGAGAGCCGCGGCGAGCGCCAGGGCA--- 364
Db 3258 AGGTTTGGGAAGAGACAACAACTCCCTCTCAGAAGCAGAGCGGATAGACAAGAACTG 3317
Qy 365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGGTGAGCATCAAGTGG 421
Db 3318 TATCTTTTAGTTTCCCTCAGATCACTTTTGGCAGCAGCCCTCTGTCACATARAAGATAG 3377
Qy 422 CGGCCAGATCAAGGAGGCCCTGCTGACACCGGCGCCGACAGCACCGTGTGGAGGAGA 481
Db 3378 GGGCCAGCTGAAGGAGGCCCTTTAGACACCGGCGCGACACACCGTGTGGAGGAGA 3437
Qy 482 TGAGCTTGGCCGCAAGTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGG 541
Db 3438 TGAACCTGGCCGCGGCGCTGAAGCCCAAGATGATCGGCGGCATCGGCGGCTTCATCAAGG 3497
Qy 542 TGCGCCAGTACGACAGATCTGTGAGATCTGCGGCAAGAGGCCATCGGCCACCTGC 601
Db 3498 TGGCCAGTAGCAGACAGTCTGTATCGAGATCTCGGCGCAACAGGCCATCGGCCACCTGC 3557
Qy 602 TGATCGGCCCAACCGGCTGAACATCATATCGGCGCGCAACATGCTGACCGAGTGGGTGCA 661
Db 3558 TGGTGGGCGCCACCGCGTGAACATCATATCGGCGCAACATGCTGACCGAGATCGGCTGCA 3617
Qy 662 CCCTGAACTTCCCATCAGCCCACTGAGACCGTGGCCGTAAGCTGAAGCCCGGCATGG 721
Db 3618 CCCTGAACTTCCCATCAGCCCACTGAGACCGTGGCCGTAAGCTGAGCCCGGCATGG 3677
Qy 722 ACGCCCAAGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTGACCGCA 781
Db 3678 ACGCCCAAGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTGTTGGAGA 3737
Qy 782 TCTGGAGGAGATGAGAGAGGCGCAGATCACCAGATCGGCCCGCAGAACCCCTACA 841
Db 3738 TCTGACCGAGATGAGAAGGAGGCGCAAGATCAGCAAGATCGGCCCGCAGAACCCCTACA 3797
Qy 842 ACACCCCGCTTGGCCATCAAGAAAGAGAGCAGCACCAGTGGCGCAAGCTGGTGGACT 901
Db 3798 ACACCCCGCTTGGCCATCAAGAAAGAGAGCAGCACCAGTGGCGCAAGCTGGTGGACT 3857
Qy 902 TCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGAGTGGGCGATCCCCAC 961
Db 3858 TCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGAGTGGGCGATCCCCAC 3917
Qy 962 CGCGCGCTGAAGAAGAGAGCGTGACCGTGTGGAGTGGCGGCGCCTACTTCA 1021
Db 3918 CGCGCGCTTGAAGCAGAAGAGAGCGTGACCGTGTGGAGTGGCGGCGCCTACTTCA 3977
Qy 1022 GCGTGGCCCTGGACGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGCAGCATCA 1081
Db 3978 GCGTGGCCCTGGACAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGCAGCATCA 4037
Qy 1082 ACGAGACCCCGGCGATCCGCTACAGTACAGTGTCTGCGGCGGCGGCTGGAGGCGAC 1141
Db 4038 ACGAGACCCCGGCGATCCGCTACAGTACAGTGTCTGCGGCGGCGGCTGGAGGCGAC 4097
Qy 1142 CCAGCATCTTCAGAGCAGATGACCAAGATCTCGAGCGCTTCGCGCGCGCAACCCG 1201
Db 4098 CGGCCATCTTCAGTGTGAGATGACCAAGATCTTCGAGCGCTTCCGCAAGCAGAACCCG 4157
Qy 1202 AGATCGTGTATACCA-----GGCCCGCCCTGTAGTGGCGGCGAGCACCTGGAGATCG 1255
Db 4158 ACATCGTGTATACCAAGTATGAGCACCTGTACGTGGCGAGCACCTGGAGATCGGCC 4217
Qy 1256 AGCACCGGCGGAGATCGAGAGCTGCGCAAGCACCTGCTGGCTGGGCTTCACCAACC 1315
Db 4218 AGCACCGCAACCAAGATCGAGGAGCTGCGCCAGCACCTGCTGGCTGGGCTTCACCAACC 4277

Qy 1316 CCGACAGAAGCACCAGAAAGGAGCCCTTCTCTGTGATGGGCTACGAGCTGCACCCG 1375
Db 4278 CCGACAGAAGCACCAGAAAGGAGCCCTTCTCTGTGATGGGCTACGAGCTGCACCCG 4337
Qy 1376 ACAAGTGAACCGTGCAGCCCATCGAGCTGGCCGAGAAAGAGAGCTGGACCGTGAACGACA 1435
Db 4338 ACAAGTGAACCGTGCAGCCCATCGTCTGCCGAGAAAGAGAGCTGGACCGTGAACGACA 4397
Qy 1436 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGGCGATCAAGGTGC 1495
Db 4398 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGGCGATCAAGGTGC 4457
Qy 1496 GCCAGCTGTGCAAGCTGCTGGCGGCGCCAGAGCCCTGACCGACATGCTGCTCCCTGACCG 1555
Db 4458 GCCAGCTGTGCAAGCTGCTGGCGGCGCCAGAGCCCTGACCGAGTGTGCTCCCTGACCG 4517
Qy 1556 AGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCAGATCTCTCGCGAGCCCGTGCACGCG 1615
Db 4518 AGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCAGATCTCTCAAGGAGCCCTGACCGCG 4577
Qy 1616 TGTACTAGACCCCGACAGAAAGACCTGGTGGCGGAGATCCAGAAAGCGGCGCACGACCA 1675
Db 4578 TGTACTAGACCCCGACAGAAAGACCTGGTGGCGGAGATCCAGAAAGCGGCGCACGACCA 4637
Qy 1676 GGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCAAGA 1735
Db 4638 GGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCAAGA 4697
Qy 1736 TGGCACCGCCCAACCAACGACGTGAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCA 1795
Db 4698 TGAAGGCGCCCAACCAACGACGTGAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCA 4757
Qy 1796 TGGAGAGCATGCTGATCTGGGGCAGAGCCCAAGTTCGCGCTGCCCATCCAGAAAGAGA 1855
Db 4758 CCGAGAGCATGCTGATCTGGGGCAGAGCCCAAGTTCGAGTCCGCTCCCATCCAGAAAGAGA 4817
Qy 1856 CCGAGAGCATGCTGAGCCGACTACTGGCAGCGCACCTGGATCCCGAGTGGGAGTTCG 1915
Db 4818 CCGAGAGCATGCTGAGCCGACTACTGGCAGCGCACCTGGATCCCGAGTGGGAGTTCG 4877
Qy 1916 TGAACACCCCGCTGCTGAGCTGAGTGTGTACCGAGTGGAGAGAGCCCATCATCGCG 1975
Db 4878 TGAACACCCCGCTGCTGAGCTGAGTGTGTACCGAGTGGAGAGAGCCCATCATCGCG 4937
Qy 1976 CCGAGACCTTCTAGTGCAGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCT 2035
Db 4938 CCGAGACCTTCTAGTGCAGCGCGCCCAACCGGAGACCAAGTGGCAAGCGCGCT 4997
Qy 2036 ACGTGACCGACCGGCGCGCGCAGAGATCGTGAAGCTGACCGAGACCAACCAAGAGAGA 2095
Db 4998 ACGTGACCGACCGGCGCGCGCAGAGAGTGGTGGCCCTGACCGACCAACCAAGAGAGA 5057
Qy 2096 CCGAGCTGACGGCCATCAGCTGGCCCTGCGAGGACAGCGGCGAGGAGTGAACATCGTGA 2155
Db 5058 CCGAGCTGACGGCCATCAGCTGGCCCTGCGAGGACAGCGGCGTGGAGTGAACATCGTGA 5117
Qy 2156 CCGAGAGCAGTACGCTTGGGCTATCCAGCGCCAGCGCCAGAGAGCGAGAGCGAGC 2215
Db 5118 CCGAGAGCAGTACGCTTGGGCTATCCAGCGCCAGCGCCAGAGAGCGAGAGCGAGC 5177
Qy 2216 TGGTGAACAGATCATCGAGAGCTGATCAAGAAAGAGAGTGTACCTGAGCTGGGTGC 2275
Db 5178 TGGTGAACAGATCATCGAGAGCTGATCAAGAAAGAGAGTGTACCTGAGCTGGGTGC 5237
Qy 2276 CCGCCCAAGAGGCGATCGCGGCAACGAGCATCGAAGCTGTGAGCAAGGCGATCC 2335
Db 5238 CCGCCCAAGAGGCGATCGCGGCAACGAGCATCGAGAGCTGTGAGCGCGGCTACCTCC 5297
Qy 2336 GCAAGTGTCTTCTTGGACGCGATCGATGGCGGCTGCTGATCTACCAAGTA 2387
Db 5298 GCAAGTGTCTTCTTGGACGCGATCGAAGAGCCCGAGGAGGAGCGAGAG 5349

Db 4041 ACAACGAGACCCCGGCATCCGCTACCAAGTACAAGCTGCTGCCCGCAGGGCTGGAAGGCA 4100
Qy 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCCTGGAGCCCTTCGCGCCCGCAACC 1198
Db 4101 GCCCGCCATCTTCCAGTGCAAGATGACCAAGATCTTGGAGCCCTTCCGCAAGCAGAAC 4160
Qy 1199 CCGAGATCGTATACCA-----GGCCCCCTGTAGTGGGAGCGACCTGGAGATCG 1252
Db 4161 CCGACATCGTATACCAAGTACATGACCAACCTGTACGTGGGAGCGACCTGGAGATCG 4220
Qy 1253 GCCAGCAGCGGCGCAAGATCAGAGATGCGCAAGCAGCTCTGCTGGGCTGGGGCTTCACCA 1312
Db 4221 GCCAGCAGCGGCGCAAGATCAGAGATGCGCAAGCAGCTCTGCTGGGCTGGGGCTTCACCA 4280
Qy 1313 CCCCAGCAAGAACACAGAGAGCGCCCTTCTCTGGATGGCTACGAGCTGCACC 1372
Db 4281 CCCCAGCAAGAACACAGAGAGCGCCCTTCTCTGGATGGCTACGAGCTGCACC 4340
Qy 1373 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCGCGGCAAGGAGAGCTGGACGTGAACG 1432
Db 4341 CCGACAAGTGACCGTGCAGCCCATCGTGTGCGCGGCAAGGAGAGCTGGACCGTGAACG 4400
Qy 1433 ACATCCAGAAGCTGTGGCAAGCTGAAGTGGGCGCAGCAGATCTACCCCGCATCAAGG 1492
Db 4401 ACATCCAGAAGCTGTGGCAAGCTGAAGTGGGCGCAGCAGATCTACCCCGCATCAAGG 4460
Qy 1493 TGCGCCAGCTGTGAAGTGTGCGCGGCGCAAGGCGCTGACCGACATCTGTCGCCCTGA 1552
Db 4461 TGCGCCAGCTGTGAAGTGTGCGCGGCGCAAGGCGCTGACCGAGTGTGTCGCCCTGA 4520
Qy 1553 CCGAGGAGCGGAGCTGGAGTGGCGGAGAACCGCGAGATCTCTGGGAGCGCCGTGCAAG 1612
Db 4521 CCGAGGAGCGGAGCTGGAGTGGCGGAGAACCGCGAGATCTCTGAAGGAGCGCCGTGCAAG 4580
Qy 1613 GCCTGTACTAGCAACCCAGCAAGCACTGTGTGGCGGAGATCCAGAAAGCGCCAGCAGC 1672
Db 4581 GCCTGTACTAGCAACCCAGCAAGCACTGTGTGGCGGAGATCCAGAAAGCGCCAGGCGC 4640
Qy 1673 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1732
Db 4641 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCC 4700
Qy 1733 AGATGGCACCAGCCACACCAACAGCTGAAGCAGCTGACCGAGGCGGTGAGAGATCG 1792
Db 4701 GCATGAAGGGCGCCACCAACAGCAGTGAAGCAGCTGACCGAGGCGGTGAGAGATCG 4760
Qy 1793 CCATGGAGAGCATCTGATCTGGGGCAAGCCCAAGTTCGGCTCGCCCATCCAGAGG 1852
Db 4761 CCACCGAGAGCATCTGATCTGGGGCAAGCCCAAGTTCAGCTGCCCATCCAGAGG 4820
Qy 1853 AGACCTGGGAGACCTGTGGACCGACTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1912
Db 4821 AGACCTGGGAGGCTGTGTGACCGAGTACTGGCAGGCGACCTGGATCCCGAGTGGGAGT 4880
Qy 1913 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGCCCATCATCG 1972
Db 4881 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGCCCATCATCG 4940
Qy 1973 GCGCGAGAGACCTTACCTGTGAGCGGCGCGCAACCGCAGACCAAGATCGGCAAGGCGC 2032
Db 4941 GCGCGAGAGACCTTACCTGTGAGCGGCGCGCAACCGCAGACCAAGATCGGCAAGGCGC 5000
Qy 2033 GCTACGTACCGACCGGGCGCGCAGAGATCGTGAAGCTGACCGAGACCAACCAAGCAGA 2092
Db 5001 GCTACGTGAACCGCGCGCGCGCAACCGCAGACCAAGCTGGCGAGGCGCG 5060
Qy 2093 AGACCGAGCTGCAGGCCCATCCAGCTGGCCCTGCAAGGAGAGCGGAGCGAGTGAACATCG 2152
Db 5061 AGACCGAGCTGCAGGCCCATCCAGCTGGCCCTGCAAGGAGAGCGGCGCGTGAACATCG 5120
Qy 2153 TGACCCGACGACAGTACGCCCTGGGCATCATCCAGGCGCGCGAGCGAGAGCGAGAGCG 2212
Db 5121 TGACCCGACGACAGTACGCCCTGGGCATCATCCAGGCGCGCGAGCGAGAGCGG 5180

Qy 2213 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 2272
Db 5181 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 5240
Qy 2273 TGCCCGCCCAACAAGGGCATCGCGGCAACGAGCAGATCGAAGCTGGTGAAGAGGCA 2332
Db 5241 TGCCCGCCCAACAAGGGCATCGCGGCAACGAGCAGTGGAGCGCTGGTGAAGCGCGCA 5300
Qy 2333 TCCGCAAGTCTGTCTCTGAGCGCATCGATGGCGGATCTGTGATCTACCAAGTA 2387
Db 5301 TCCGCAAGTCTGTCTCTGAGCGCATCGAAGCGCCAGGAGGAGCAGGAA 5355

RESULT 13
AX427935
LOCUS AX427935 9782 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 173 from Patent WO0232943.
ACCESSION AX427935
VERSION AX427935.1 GI:21538022
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 173 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
source 1..9782
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2339 a 2895 c 2719 g 1829 t
ORIGIN

Query Match 80.9%; Score 1993.4; DB 6; Length 9782;
Best Local Similarity 91.18; Pred. No. 2.1e-200;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;

Qy 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCAACATCTGTATGAGGCGAGCAACT 70
Db 3579 TGGCCGAGGCGCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGGCAACT 3638
Qy 71 TCAAGGCGCCCAACGGCGATCATCAAGTCTTCAACTGGCGCAAGAGGCGCACATCGGCC 130
Db 3639 TCCGCAACCCAGCGCAAGATCGTGAAGTCTTCAACTGGCGCAAGAGGCGCACACCGCC 3698
Qy 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGAAGAGGCGCCACACAGA 190
Db 3699 GCAACTGCGCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGAAGAGGCGCCACACAGA 3758
Qy 191 TGAAGGACTGCACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGACCTGCGCTTCCCGCAGG 250
Db 3759 TGAAGGACTGCACCGAGCGACAGCGCTAA-----TAGGGAAGATCTGCGCTTCCACAG 3812
Qy 251 GCAAGGCGCGCGAGTCTCCCGAGCAGAGAGACCGCGCCCAACAGCCCCACAGCGCGGAGC 310
Db 3813 GGAAGGCGCGGGAATTTTCTTTCAGAGCAGACAGACAGCGCCACAGCGCCACAGAGAGAGC 3872
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGGCA 364
Db 3873 TTCAGGTTTGGGGAAGAGAGACAACACTCCCTCTCAGAGCAGGAGGCCCATAGCAAGGAA 3932
Qy 365 -----CCCTGAACCTTCCCCAGATCACTCTGTGGCAGCGCCCTTGTGTGAGCATCAAG 418
Db 3933 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGTGGCAGCGACCCCTCTCTCAATAAAGA 3992
Qy 419 TGGCGCGCCAGATCAAGAGGCGCCCTGTGTGACACCGGGCGCGCAGACACCGTGTGGAGG 478
Db 3993 TAGGGGCGCAGCTGAAGGAGGCGCCCTTCTAGACCGCGCGCGAGCACACCGTGTGGAGG 4052

FEATURES		Location/Qualifiers	
source		1..9783	
		/organism="synthetic construct"	
		/db_xref="taxon:32630"	
		/note="Plasmid pVR1012x/s containing HIV genes"	
BASE COUNT	2339 a	2897 c	2723 g 1824 t
ORIGIN			
Query Match 80.9%; Score 1993.4; DB 6; Length 9783;			
Best Local Similarity 91.1%; Pred. No. 2.1e-200;			
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;			
QY	14	TGCGCGAGCCATGAGCCAGG---CCACAGGCGCAACATCTGTATGAGCGCAGCAACT	70
Db	2964	TGCGCGAGCCATGAGCCAGGTCACCAACAGCCACCATATGATGAGCGCGCAACT	3023
QY	71	TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGGCC	130
Db	3024	TCCGCAACCAAGCGCAAGATCGTGAAGTGTCTCAACTGCGGCAAGGAGGGCCACCGGCC	3083
QY	131	GCAACTGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA	190
Db	3084	GCAACTGCCCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA	3143
QY	191	TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACCTGGCTTCCCCAGG	250
Db	3144	TGAAGGACTGCACCGAGCGACAGCTAA-----TAGGGAAGATCTGGCCCTTCCCAACAG	3197
QY	251	GCAAGGCCCGGAGTTCCCGAGCGAGGACAGCCGCGCCCAACAGCCCCACACGCGCGAGC	310
Db	3198	GGAGGGCCAGGGAATTTCTTTCAGAGCAGACACAGAGCCCAACAGCCCCACCAAGAGAGC	3257
QY	311	TGCAGTGTGCGGG-----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCCAGGGCA	364
Db	3258	TTCAGGTTTGGGAGAGACAACTCCCTCTCAGAGCAGGAGCGCGATAGACAAGAA	3317
QY	365	-----CCCTGAACTTCCCGCAGATCACCTGTGTGCGAGCGCCCTTGGTGAAGCATCAGG	418
Db	3318	CTGTATCTTTAGCTTTCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAATAAAGA	3377
QY	419	TGGCGGCCAGATCAAGGAGCCCTGTCTGGACACCGCGCGCGACACACCGTCTGGAGG	478
Db	3378	TAGGGGCCAGCTGAAGGAGGCGCCCTTCTAGACACCGCGCGCGACACACCGTCTGGAGG	3437
QY	479	AGATGAGCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCA	538
Db	3438	AGATGAACTTCCCGGCGCGCTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCA	3497
QY	539	AGGTGCGCCAGTAGACACAGATCTCTGTATCGAGATCTCGGCAAGAGGCCATCGGCACCG	598
Db	3498	AGGTGGGCGAGTACCAACAGATCTCTGTATCGAGATCTGCGGCGCACCAAGGCGCATCGCACCG	3557
QY	599	TGCTGATCGGGCCCCACCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGTGGGCT	658
Db	3558	TGCTGGTGGGGCCCCACCCCGTGAACATCATCGGCCCGCAACCTGCTGACCCAGATCGGCT	3617
QY	659	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCA	718
Db	3618	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCA	3677
QY	719	TGGAGGCCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG	778
Db	3678	TGGAGGCCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTTGGTGG	3737
QY	779	CCATCTCGGAGGAGATGGAGAGGAGGCGCAAGATCACCAAGATCGGCCCGGAGAACCCCT	838
Db	3738	AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAGCAAGATCGGCCCGGAGAACCCCT	3797
QY	839	ACAAACCCCGCTGTTCCGCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGGTGG	898
Db	3798	ACAAACCCCGCTGTTCCGCATCAAGAAAGAGGACACCAAGTGGCGCAAGCTGGTGG	3857
QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGCAGTGGGATCCCCC	958
Db	3858	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGGAGGTGCAGTGGGATCCCCC	3917
QY	959	ACCCCGCGCGCTTGAAGAAGAGAGCGGTGACCGTGTGAGAGTGGGCGACCGCCTACT	1018
Db	3918	ACCCCGCGCGCTTGAAGCAGAGAGAGCGGTGACCGTGTGAGAGTGGGCGACCGCCTACT	3977
QY	1019	TCAGCGTCCCGCTTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCA	1078
Db	3978	TCAGCGTCCCGCTTGGACAAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCA	4037
QY	1079	ACAAAGAGACCCCGGCATCCGCTACCAAGTACAAAGTGTCTGCCCGAGGGCTTGAAGGGCA	1138
Db	4038	ACAAAGAGACCCCGGCATCCGCTACCAAGTACAAAGTGTCTGCCCGAGGGCTTGAAGGGCA	4097
QY	1139	GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTTCGCGCCCGCAACC	1198
Db	4098	GCCCCAGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTTCGCGCAAGCAGAAC	4157
QY	1199	CCGAGATCGTGATCTACCA-----GGCCCGCTGTAGTGGGAGGACCTGGAGATCG	1252
Db	4158	CCGACATCTGTATCTACCAAGTACATGGACACCTGTAGTGGGAGGACCTGGAGATCG	4217
QY	1253	GCCAGCACCGCGCAAGATCGAGGAGCTTGCCAAAGCAGCTCTGCGTGGGCTTTCACCA	1312
Db	4218	GCCAGCACCGCGCAAGATCGAGGAGCTTGCCAAAGCAGCTCTGCGTGGGCTTTCACCA	4277
QY	1313	CCCCGACAAAGAACACCAAGAGGAGCCCCCTTCTCTGTGTGATGGGCTACGAGCTGCACC	1372
Db	4278	CCCCGACAAAGAACACCAAGAGGAGCCCCCTTCTCTGTGTGATGGGCTACGAGCTGCACC	4337
QY	1373	CCGACAAAGTGCACCGTGCAGCCCATCGAGCTGCCGAGAGGAGAGCTTGAACCTGAAACG	1432
Db	4338	CCGACAAAGTGCACCGTGCAGCCCATCGTGTGCCGAGAGGACAGCTGGAACCTGAAACG	4397
QY	1433	ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGACGAGATCTACCCCGGCAATCAAGG	1492
Db	4398	ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGACGAGATCTACCCCGGCAATCAAGG	4457
QY	1493	TGCGCGAGCTGTCAAGCTGTCTGCGCGGCGCAAGGCGCTTACCGCATCTGTCGCCCTGA	1552
Db	4458	TGCGCGAGCTGTCAAGCTGTCTGCGCGGCGCAAGGCGCTTACCGAGGCTGTGTCGCCCTGA	4517
QY	1553	CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGCGCGAGCCCGTGCACG	1612
Db	4518	CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAAGGAGCCCGTGCACG	4577
QY	1613	GGGTGTACTAGACCCCGACGAAAGCACTGGTGGCCGAGATCCAGAAGCAGGCGCACGACC	1672
Db	4578	GGGTGTACTAGACCCCGACGAAAGCACTGTATCGCCGAGATCCAGAAGCAGGCGCACGACC	4637
QY	1673	AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAACCTTGAAGAACCGGCAAGTACGCCA	1732
Db	4638	AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAACCTTGAAGAACCGGCAAGTACGCCA	4697
QY	1733	AGATGCGCACCGCCCGACACCAACGACGTGAAGCAGCTGACCGAGGCGCTGCGAGAGATCG	1792
Db	4698	GCATGAAGGCGCGCCACACCAACGACGTGAAGCAGCTGACCGAGGCGCTGCGAGAGATCG	4757
QY	1793	CCATGGAGAGCATCGTGTATCTTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGG	1852
Db	4758	CCACCGAGAGCATCGTGTATCTTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGG	4817
QY	1853	AGACCTGGGAGACCTGGTGGACCGGACTTCTGGCAGGCCACCTGGATCCCGAGTGGGAGT	1912
Db	4818	AGACCTGGGAGGCGCTGGTGGACCGGAGTACTGGCAGGCGCACTGGATCCCGAGTGGGAGT	4877
QY	1913	TGCTGAACACCCCGCTTGTGTGAAGCTGTGTACACGCTGGAAGAGGAGGAGCCCATCATCG	1972
Db	4878	TCGTGAACACCCCGCTTGTGTGAAGCTGTGTACCAAGCTTGGAGAGGAGGAGCCCATCATCG	4937
QY	1973	GGCGCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG	2032

```
Db 4938 GCGCCGAGACCTTCTACGTGGACGGCGCGCCCAACCCGCGAGACCAAGCTGGGCAAGGCCG 4997
QY 2033 GCTACGTACCGACCGGGCGGCGGAGAGATCGTGAAGCTGACCGACAGACCAACACAGA 2092
Db 4998 GCTACGTACCGACCGGGCGGCGGAGAGTGGTGGCCCTGACCGACAGACCAACACAGA 5057
QY 2093 AGACCGAGTGCAGGCGCATCCAGTGGCCCTGCGAGGACAGCGGAGGTGAACATCG 2152
Db 5058 AGACCGAGTGCAGGCGCATCCAGTGGCCCTGCGAGGACAGCGGAGGTGAACATCG 5117
QY 2153 TGACCGGACAGCGAGTACGCGCTGGGCATCTCCAGGCGGCGGAGGCGGAGCG 2212
Db 5118 TGACCGGACAGCGAGTACGCGCTGGGCATCTCCAGGCGGCGGAGGCGGAGCG 5177
QY 2213 AGTGTGTGAACAGCATCGAGCAGCTGATCAAGAAGGAGAGGTGACTGAGCTGGG 2272
Db 5178 AGTGTGTGAGCAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGACTGAGCTGGG 5237
QY 2273 TGCCCGCCACAGGGGATCGGGGCGCAACGAGCAGATCGACAAAGCTGGTGAAGCAAGGCA 2332
Db 5238 TGCCCGCCACAGGGGATCGGGGCGCAACGAGCAGGTGGACGGCTGGTGAAGCGCCGCA 5297
QY 2333 TCCGCAAGGTGCTGTTCTTCCAGGCGCATCGATGCGGCGCATCGTATACAGTA 2387
Db 5298 TCCGCAAGGTGCTGTTCTTCCAGGCGCATCGACAAAGGCCCGAGGAGCAGAGAA 5352

RESULT 15
AX427932
LOCUS AX427932 9792 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 170 from Patent W00232943.
ACCESSION AX427932
VERSION AX427932.1 GI:21538019
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 170 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source
1. .9792
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2339 a 2898 c 2725 g 1830 t
ORIGIN

Query Match 80.9%; Score 1993.4; DB 6; Length 9792;
Best Local Similarity 91.1%; Pred. No. 2,1e-200;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;

QY 14 TGCGCGAGGCCATGAGCCAGG----CCACAGCGCCACATCTCGATGACGCGCAGCAACT 70
Db 2967 TGCGCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGACGCGGCAACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTTCACATGCGGCAAGGAGGCGGCACATCGCCC 130
Db 3027 TCCGCAACAGCGCAGATCGTGAAGTGTTCACATGCGGCAAGGAGGCGGCACACCGCCC 3086
QY 131 GCAACTGCCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGCGGCACAG 190
Db 3087 GCAACTGCCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGGCGCAAGGAGGCGGCACAG 3146
QY 191 TGAAGACTGACCGAGCGCCAGGCGCAACTTCTTCGCGAGGACCTTGGCCCTTCCCGCAGG 250
Db 3147 TGAAGACTGACCGAGCGCCAGGCGCAACTTCTTCGCGAGGACCTTGGCCCTTCCCGCAGG 3200
QY 251 GCAAGGCCCGCGAGTTCCCGAGGCGAGCAGAACCGCGCCACACCGCCAGCGCGAGC 310
```

```
Db 3201 GGAAGCGCAGGGAATTTTCTTACAGAGCAGACCAAGAGCCCAACAGCCCAACCAAGAGAGC 3260
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGGCGCGCAGCGCAGAGCA 364
Db 3261 TTCAGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 3320
QY 365 -----CCCTGAACTTCCCCCAGATCACCTCTGTGGCAGCGCCCTGTGTGAGCATCAAG 418
Db 3321 CTGTATCTTCTTAGCTTCCCTCAGATCACTCTTTGCGAGCGACCCCTCGTCACAATAAGA 3380
QY 419 TGGCGGCGCAGATCAAGAGAGCCCTGTGTGACACCGGCGCGCAGACACCTGTGTGAGG 478
Db 3381 TAGGGGCGCAGCTGAAGAGGCGCCCTTCTAGACACCGGCGCGCAGACACCTGTGTGAGG 3440
QY 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 538
Db 3441 AGATGAACTTGCCCGGCGCTGGAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 3500
QY 539 AGGTGCGCCAGTAGACGACAGATCCTGTATCGAGATCTTGGGCAAGAAGGCCATCGGCACCG 598
Db 3501 AGGTGGGCCAGTAGACGACAGATCCTGTATCGAGATCTTGGGCAAGAAGGCCATCGGCACCG 3560
QY 599 TGCTGATCGGCGCCACCGCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3561 TGCTGTTGGCGCCCGACCGCGTGAACATCATCGGCGCGCAACCTGCTGACCCAGATCGGCT 3620
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAAGTGAAGTGAAGCCCGCA 718
Db 3621 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAAGTGAAGTGAAGCCCGCA 3680
QY 719 TGACGCGCCCAAGGTGAAGCAGTGGCCCTGTGACCGGAGAGAGATCAAGGCCCTTGACCG 778
Db 3681 TGACGCGCCCAAGGTGAAGCAGTGGCCCTGTGACCGGAGAGAGATCAAGGCCCTTGAGTG 3740
QY 779 CCATCTCGGAGGAGATGGAAAGGAGGCAAGATCAACCAAGATCGGCGCGCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAAAGGAGGCAAGATCAACCAAGATCGGCGCGCGAGAACCCCT 3800
QY 839 ACAACACCCCGCTGTTCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
Db 3801 ACAACACCCCGCTGTTCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 3860
QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCAATCCCC 958
Db 3861 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCAATCCCC 3920
QY 959 ACCCGCGCGCGCTGAAGAAAGAGAGCGTGCCTGGAGGTGCTGGGAGCGGCGGACCGCTACT 1018
Db 3921 ACCCGCGCGCGCTGAAGAAAGAGAGCGTGCCTGGAGGTGCTGGGAGCGGCGGACCGCTACT 3980
QY 1019 TCAGCGTGCCTCTGGAGCAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
Db 3981 TCAGCGTGCCTCTGGAGCAGGACTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 4040
QY 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACAAAGTGTGCTGCCCCAGGGGTGGAAGGCA 1138
Db 4041 ACAACGAGACCCCGGATCCGCTACCAAGTACAAAGTGTGCTGCCCCAGGGGTGGAAGGCA 4100
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCGCTTCCGCGCGCGCAACC 1198
Db 4101 GCGCCAGCATCTTCCAGTGCAGCATGACCAAGATCTTGGAGCGCTTCCGCGCGCGCAACC 4160
QY 1199 CCAGATCTGTGATCTACCA-----GGCCCCCTGTACGTGGGCGAGGAGCTGGAGATCG 1252
Db 4161 CCAGATCTGTGATCTACCAAGTACATGACCACTGTACGTGGGCGAGGAGCTGGAGATCG 4220
QY 1253 GCAGACACCGCGCAGATCGAGAGCTGCGCAAGCAACCTGCTGCGGTGGGGGTTCACCA 1312
Db 4221 GCAGACACCGCGCAGATCGAGAGCTGCGCAAGCAACCTGCTGCGGTGGGGGTTCACCA 4280
QY 1313 CCCCCGACAGAGAGCAGACCAAGAGCGCCCTTCCCTGTGTGATGGGTAGGAGTGCACC 1372
Db 4281 CCCCCGACAGAGAGCAGACCAAGAGCGCCCTTCCCTGTGTGATGGGTAGGAGTGCACC 4340
```

QY 1373 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAGAGAGCTGGACCGTGAACG 1432
|||||
Db 4341 CCGACAAGTGGACCGTGCAGCCCATCGTCTGCCGAGAAGAGAGCTGGACCGTGAACG 4400
QY 1433 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGGCATCAAGG 1492
|||||
Db 4401 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGGCATCAAGG 4460
QY 1493 TCGCGCAGCTGTGCAAGCTGTGGCGCGCCCAAGGCCCTCAACGAGATCGTGCCCTGTA 1552
|||||
Db 4461 TCGCGCAGCTGTGCAAGCTGTGGCGCGCCCAAGGCCCTCAACGAGATCGTGCCCTGTA 4520
QY 1553 CCGAGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGGCGAGGCCCGCTGCACG 1612
|||||
Db 4521 CCGAGAGGCCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTGAAGNAGCCCGTGCACG 4580
QY 1613 GCGTGTACTAGACCCAGCAAGACCTGTGGTGGCGAGATCCAGAAAGAGGCCACGACC 1672
|||||
Db 4581 GCGTGTACTAGACCCAGCAAGACCTGTATCGCCGAGATCCAGAAAGAGGCCACGACC 4640
QY 1673 AGTGGACCTACGAGATCTACCGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1732
|||||
Db 4641 AGTGGACCTACGAGATCTACCGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 4700
QY 1733 AGATGCGCACCGCCACACCAACGACCTGAAGCAGCTGACCGAGGCCGCTGCAGAAGATCG 1792
|||
Db 4701 GCATGAAGGGCGCCACACCAACGACCTGAAGCAGCTGACCGAGGCCGCTGCAGAAGATCG 4760
QY 1793 CCATGGAGAGCATCGTATCTGGGGCAGACCCGCCAGTTCCGCCCTGCCATCCAGAAAG 1852
|||||
Db 4761 CCACGGAGAGCATCGTATCTGGGGCAGACCCGCCAGTTCAAGCTGCCATCCAGAAAG 4820
QY 1853 AGACTGGGAGACCTGTGGTGGACCTACTGGCAGGCCACCTGATCCCGAGTGGGAGT 1912
|||||
Db 4821 AGACTGGGAGACCTGTGGTGGACCTACTGGCAGGCCACCTGATCCCGAGTGGGAGT 4880
QY 1913 TCGTGAACACCCCCCTGTGAAGCTGTGTACCACTGGAGAGAGGCCCATCATCG 1972
|||||
Db 4881 TCGTGAACACCCCCCTGTGAAGCTGTGTACCACTGGAGAGAGGCCCATCATCG 4940
QY 1973 GCGCGGAGACCTTACGTGGACGGCCCGCCGACACCGCGAGACCAAGATCGGCAAGGCCG 2032
|||||
Db 4941 GCGCGGAGACCTTACGTGGACGGCCCGCCGACACCGCGAGACCAAGATCGGCAAGGCCG 5000
QY 2033 GCTACGTGACCGCGGGCGGCGAGAGATCGTGAGCCCTGACGAGACCAACCAACGAGA 2092
|||||
Db 5001 GCTACGTGACCGCGGGCGGCGAGAGTGTGCCCCCTGACCGACCAACCAACGAGA 5060
QY 2093 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCG 2152
|||||
Db 5061 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCTGGAGGTGAACATCG 5120
QY 2153 TGACCGAGACCCAGTACGCCCTGGGCATCATCCAGGCCCGACCGCGAGAGCGAGAGCG 2212
|||||
Db 5121 TGACCGAGACCCAGTACGCCCTGGGCATCATCCAGGCCCGACCGCGAGAGCGAGAGCG 5180
QY 2213 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGCTACCTGAGCTGGG 2272
|||||
Db 5181 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGCTACCTGAGCTGGG 5240
QY 2273 TGCCCGCCCAACAGGGCATCGGGGCAACGAGCAGATCGACAAGCTGGTGAAGCAAGGGCA 2332
|||||
Db 5241 TGCCCGCCCAACAGGGCATCGGGGCAACGAGCAGGTGGAGCGCCCTGGTGAAGCGCGCA 5300
QY 2333 TCCGCAAGGTGCTGTTCTCGAGGGCATCGATGGCGGCATCGTGTATCTACCACTA 2387
|||||
Db 5301 TCCGCAAGGTGCTGTTCTCGAGGGCATCGACAAGGCCCGAGAGGACGAGAA 5355

Result No.	Score	Query Match	Length	DB	ID	Description
1	2463	100.0	2463	24	ABL39960	Synthetic construc
2	2442.2	99.2	2469	24	ABL39959	Synthetic construc
3	2436.2	98.9	2457	24	ABL39961	Synthetic construc
4	2046	83.1	2306	21	AAA70479	HIV FS(-)_ProtMod
5	2027.6	82.3	9166	24	ABK91616	Modified HIV prote
6	2025.2	82.2	2312	21	AAA70481	HIV FS(-)_ProtMod
7	2019.2	82.0	2300	21	AAA70480	HIV FS(-)_ProtMod
8	2019.2	82.0	9788	24	ABK91622	Modified HIV prote
9	2005.2	81.4	9169	24	ABK91617	Modified HIV prote

PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy -
XX
XX Claim 1; Fig 9; 233pp; English.
XX
CC The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding gag, pol, vif, vpr, tat, rev, vpu, env or
CC nef (1). (i) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABB06204 to ABB06215 represent sequences used in the
CC exemplification of the present invention.
XX
SQ Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 other;

Query Match 100.0%; Score 2463; DB 24; Length 2463;
Best Local Similarity 100.0%; Pred. No. 1.4e-300;
Matches 2463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTCAGCGCCACATGCGCGAGGCCATGAGCGCCAGCGCCACACATCCTGATGAG 60
DB 1 GTCAGCGCCACATGCGCGAGGCCATGAGCGCCAGCGCCACACATCCTGATGAG 60

QY 61 GCGAGCACTTCAAGGGCCCAAGCGGATCATCAAGTCTTCACTGGGCAAGAGGGC 120
DB 61 GCGAGCACTTCAAGGGCCCAAGCGGATCATCAAGTCTTCACTGGGCAAGAGGGC 120

QY 121 CACATCGCCCACTGCGCGCCCGCCGCAAGAGGCGCTGCTGAAGTGGCGCAAGGAG 180
DB 121 CACATCGCCCACTGCGCGCCCGCCGCAAGAGGCGCTGCTGAAGTGGCGCAAGGAG 180

QY 181 GCGCACCAGATGAAGGACTGCACGAGCGCGCAGCGCCCAACTTTCTTCGCGAGGAGCTGGCC 240
DB 181 GCGCACCAGATGAAGGACTGCACGAGCGCGCAGCGCCCAACTTTCTTCGCGAGGAGCTGGCC 240

QY 241 TTCCCGCAGGCAAGGCCCGGAGTTCCTCCAGCGAGAGACCGCGCCCAAGCCCGCCACC 300
DB 241 TTCCCGCAGGCAAGGCCCGGAGTTCCTCCAGCGAGAGACCGCGCCCAAGCCCGCCACC 300

QY 301 AGCGCGAGCTGCAGGTGCGCGGGGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCCAG 360
DB 301 AGCGCGAGCTGCAGGTGCGCGGGGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCCAG 360

QY 361 GCGACCTTGAATCTCCCGCAGATCACCTGTGGAGCGCCCGCTGGTGAATCAAGTGTG 420
DB 361 GCGACCTTGAATCTCCCGCAGATCACCTGTGGAGCGCCCGCTGGTGAATCAAGTGTG 420

QY 421 GCGCGCAGATCAAGGAGCGCTGCTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 480
DB 421 GCGCGCAGATCAAGGAGCGCTGCTGACACCGCGCGCGCGCGCGCGCGCGCGCGCGAG 480

QY 481 ATGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAG 540
DB 481 ATGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAG 540

QY 541 GTGCGCAGATAGGACAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 541 GTGCGCAGATAGGACAGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 601 CTGATCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCTGC 660
DB 601 CTGATCGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCTGC 660

QY 661 ACCCTGAATCTCCCGCATCAGCGCCCATCGAGACCGCTGCGCGTGAAGCTGAAGCCCGGCTATG 720
DB 661 ACCCTGAATCTCCCGCATCAGCGCCCATCGAGACCGCTGCGCGTGAAGCTGAAGCCCGGCTATG 720

QY 721 GACGGCCCCAAGGTGAAGCACTGGCCCTTCACGAGGAGAGATCAAGGCCCTGACCGCC 780

DB 721 GACGGCCCCAAGGTGAAGCACTGGCCCTTCACGAGGAGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGGAGAGATGGAGAAGGAGGCAAGATCACCAGATCGCCCGGAGAACCCCTTAC 840
DB 781 ATCTGGAGAGATGGAGAAGGAGGCAAGATCACCAGATCGCCCGGAGAACCCCTTAC 840

QY 841 AACACCCCTGTTTCGCCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCTGTTTCGCCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGGAC 900

QY 901 TTCCGCGAGGTGAACAAAGCGCACCCAGACTTCTGGAGGTGACAGTGGGATCCCCAC 960
DB 901 TTCCGCGAGGTGAACAAAGCGCACCCAGACTTCTGGAGGTGACAGTGGGATCCCCAC 960

QY 961 CCAGCGCGCTGGAAGAAGAAGAGCGTACCTGCTGGACGTGGGCGAGCGCTACTTTC 1020
DB 961 CCAGCGCGCTGGAAGAAGAAGAGCGTACCTGCTGGACGTGGGCGAGCGCTACTTTC 1020

QY 1021 AGCTGCGCTGGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCCAGCATCAAC 1080
DB 1021 AGCTGCGCTGGACGAGGACTTCCGCAAGTACACCGCTTTCACCATCCCCAGCATCAAC 1080

QY 1081 AACGAGACCCCGCATCCGCTACAGTACAACTGACAGTGGCGGAGGCTGGAAGGCGAGC 1140
DB 1081 AACGAGACCCCGCATCCGCTACAGTACAACTGACAGTGGCGGAGGCTGGAAGGCGAGC 1140

QY 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCGCTTCCGCGCGCCCAACCCC 1200
DB 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGAGCGCTTCCGCGCGCCCAACCCC 1200

QY 1201 GAGATCGTGTATCTACAGGCCCCCTGTACGTGGGACGAGCTGGAGATCGGCCAGCAC 1260
DB 1201 GAGATCGTGTATCTACAGGCCCCCTGTGTGTGGGACGAGCTGGAGATCGGCCAGCAC 1260

QY 1261 CGCGCAGATCGAGGAGCTGCGCAAGCAGCTGTGCTGGGCTTCCACACCCCGCAC 1320
DB 1261 CGCGCAGATCGAGGAGCTGCGCAAGCAGCTGTGCTGGGCTTCCACACCCCGCAC 1320

QY 1321 AAGAAGCACCAGAGAGGAGCCCCCTTCTGTGTGGATGGGCTACGAGCTGCACCCCGCACAG 1380
DB 1321 AAGAAGCACCAGAGAGGAGCCCCCTTCTGTGTGGATGGGCTACGAGCTGCACCCCGCACAG 1380

QY 1381 TGGACCTGAGCGCCATCGAGCTGCGCGAGAGAGAGCTGGACCTGACACACATCCAG 1440
DB 1381 TGGACCTGAGCGCCATCGAGCTGCGCGAGAGAGAGCTGGACCTGACACACATCCAG 1440

QY 1441 AAGCTGTGGCAAGCTGAAGTGGCGCAGCAGATCTACCCCGCATCAAGTGGCGCAG 1500
DB 1441 AAGCTGTGGCAAGCTGAAGTGGCGCAGCAGATCTACCCCGCATCAAGTGGCGCAG 1500

QY 1501 CTGTGCAAGCTGCTGCGCGGCGCCAAAGCGCTGACCGACATCGTCCCTGACCGAGGAG 1560
DB 1501 CTGTGCAAGCTGCTGCGCGGCGCCAAAGCGCTGACCGACATCGTCCCTGACCGAGGAG 1560

QY 1561 GCGAGCTGGAGCTGCGCGGAGAACCGGAGATCTGCGGAGCGCTGACCGCGCTGTAC 1620
DB 1561 GCGAGCTGGAGCTGCGCGGAGAACCGGAGATCTGCGGAGCGCTGACCGCGCTGTAC 1620

QY 1621 TACGACCCAGCAGCAGAGCTGCTGGCGCGAGATCCAGAGCAGGCGCAGACAGTGGAGC 1680
DB 1621 TACGACCCAGCAGCAGAGCTGCTGGCGCGAGATCCAGAGCAGGCGCAGACAGTGGAGC 1680

QY 1681 TACAGATCTACCGAGGCGCTTCAAGACCTTGAAGACCGGCAAGTACCGCAAGTGGCG 1740
DB 1681 TACAGATCTACCGAGGCGCTTCAAGACCTTGAAGACCGGCAAGTACCGCAAGTGGCG 1740

QY 1741 ACCGCGCCACACCAAGCAGCTGAAGCAGCTGACGAGCGCTGCAGAAAGTCCGCATGGAG 1800
DB 1741 ACCGCGCCACACCAAGCAGCTGAAGCAGCTGACGAGCGCTGCAGAAAGTCCGCATGGAG 1800

QY 1801 AGCATCGTGTATGGGCGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGGAGAGCTGG 1860


```
QY 661 ACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGAGCTGAAGCCCGCGCATG 720
Db 661 ACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGAGCTGAAGCCCGCGCATG 720
QY 721 GAGCGCCCAAGGTGAAGCAGTGGCCCTGACGAGGAGAGATCAAGGCCCTGACCGCC 780
Db 721 GAGCGCCCAAGGTGAAGCAGTGGCCCTGACGAGGAGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGGAGAGATGGAGAGAGGAGGCAAGATCACCAGATCGGCCCGCGAGAACCCCTAC 840
Db 781 ATCTGGAGAGATGGAGAGAGGAGGCAAGATCACCAGATCGGCCCGCGAGAACCCCTAC 840
QY 841 AACACCCCTGTTGCGCATCAAGAAAGAGACAGCAAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCTGTTGCGCATCAAGAAAGAGAGACAGCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGGAGGTGAACAAAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCAC 960
Db 901 TTCGGAGGTGAACAAAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCAC 960
QY 961 CCGCGGCGCTGAGAGAGAGAGAGCGTGCCTGCTGACGTGGCGAGCGCTACTTC 1020
Db 961 CCGCGGCGCTGAGAGAGAGAGAGCGTGCCTGCTGACGTGGCGAGCGCTACTTC 1020
QY 1021 AGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGCATCAAC 1080
Db 1021 AGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGCATCAAC 1080
QY 1081 AAGGAGACCCCGCGATCPCGCTACCAAGTACAGCTGCTGCCCGAGGCTGGAAGGGCAGC 1140
Db 1081 AAGGAGACCCCGCGATCPCGCTACCAAGTACAGCTGCTGCCCGAGGCTGGAAGGGCAGC 1140
QY 1141 CCCAGCATCTCCAGACGATGACCAAGATCCTGAGCGCTTCCGCGCGCGCAACCCC 1200
Db 1141 CCCAGCATCTCCAGACGATGACCAAGATCCTGAGCGCTTCCGCGCGCGCAACCCC 1200
QY 1201 GAGATCGTGTATACCA - - - - -GGCCCGCTGTACCTGGCGAGCGACTGGAGATCGGC 1254
Db 1201 GAGATCGTGTATACCAAGTACATGGACGACTGTACCTGGCGAGCGACTGGAGATCGGC 1254
QY 1255 CAGACCGCGCAAGATCGAGGAGCTGCGAAGACCTGTGCGCTGGGCTTACCAACC 1314
Db 1255 CAGACCGCGCAAGATCGAGGAGCTGCGAAGACCTGTGCGCTGGGCTTACCAACC 1314
QY 1315 CCGACAAGAGACCAAGAGAGCGCCCTTCTGCTGGATGGCTACGAGCTGCACCCC 1374
Db 1315 CCGACAAGAGACCAAGAGAGCGCCCTTCTGCTGGATGGCTACGAGCTGCACCCC 1374
QY 1375 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACGAC 1434
Db 1375 GACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACGAC 1434
QY 1435 ATCCAGAAGCTGTGGGCAAGCTGAATCGGCGCAGCCAGATCTACCCCGGATCAAGGTG 1494
Db 1435 ATCCAGAAGCTGTGGGCAAGCTGAATCGGCGCAGCCAGATCTACCCCGGATCAAGGTG 1494
QY 1495 CGCCAGCTGTGAAGCTGTGCGCGCGCGCAAGCCCTGACCGACATCTGCGCCCTGACC 1554
Db 1495 CGCCAGCTGTGAAGCTGTGCGCGCGCGCAAGCCCTGACCGACATCTGCGCCCTGACC 1554
QY 1555 GAGAGGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGCGAGCCCGTGCACGGC 1614
Db 1555 GAGAGGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGCGAGCCCGTGCACGGC 1614
QY 1615 GTGTACTACGACCCCAAGGAGCTGGTGGCGGAGATCAGAAAGAGGCGCAGCAG 1674
Db 1615 GTGTACTACGACCCCAAGGAGCTGGTGGCGGAGATCAGAAAGAGGCGCAGCAG 1674
QY 1675 TGGACCTACGAGATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCGAAG 1734
Db 1675 TGGACCTACGAGATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCGAAG 1734
QY 1735 ATGCGACCCCGCACACCAAGCAGCTGAAGCAGCTGACCGAGCGCTGCAGAAAGATCGCC 1794
Db 1735 ATGCGACCCCGCACACCAAGCAGCTGAAGCAGCTGACCGAGCGCTGCAGAAAGATCGCC 1794
QY 1795 ATGAGAGACATCGTGTCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAAGAG 1854
Db 1795 ATGAGAGACATCGTGTCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAAGAG 1854
QY 1855 ACCTGGAGACCTGGTGGACCGCTACTGTCAGCGCCACCTGGATCCCGGAGTGGAGTTC 1914
Db 1855 ACCTGGAGACCTGGTGGACCGCTACTGTCAGCGCCACCTGGATCCCGGAGTGGAGTTC 1914
QY 1915 GTGAACACCCCGCTGCTGAGCTGTGTTACAGCTGGAGAGAGAGCCCATCATCGGC 1974
Db 1915 GTGAACACCCCGCTGCTGAGCTGTGTTACAGCTGGAGAGAGAGCCCATCATCGGC 1974
QY 1975 GCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCAGACCAAGATCGGCAAGCGCGGC 2034
Db 1975 GCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCAGACCAAGATCGGCAAGCGCGGC 2034
QY 2035 TAGCTGACCGACCGCGCGCGCAAGATCGTGAAGCTGACCGAGACCAACCAAGCAAG 2094
Db 2035 TAGCTGACCGACCGCGCGCGCAAGATCGTGAAGCTGACCGAGACCAACCAAGCAAG 2094
QY 2095 ACCGAGCTGACGCGCATCCAGCTGGCGCTGCAGGACAGCGGAGAGTGAACATCGTG 2154
Db 2095 ACCGAGCTGACGCGCATCCAGCTGGCGCTGCAGGACAGCGGAGAGTGAACATCGTG 2154
QY 2155 ACCGACAGCAGTACGCGCTGGCGCATCATCCAGCGCCAGCGCCGAGAGCGAGCGAG 2214
Db 2155 ACCGACAGCAGTACGCGCTGGCGCATCATCCAGCGCCAGCGCCGAGAGCGAGCGAG 2214
QY 2215 CTGCTGAACAGATCATCGAGCAGCTCATCAAGAGAGAGAGTGTACCTGAGCTGGGTG 2274
Db 2215 CTGCTGAACAGATCATCGAGCAGCTCATCAAGAGAGAGAGTGTACCTGAGCTGGGTG 2274
QY 2275 CCGCCCGACAAAGGCGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAAGAGGCGATC 2334
Db 2275 CCGCCCGACAAAGGCGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAAGAGGCGATC 2334
QY 2335 CGCAAGTGTCTTCTGGAGCGCATCGATGGCGCATCGATGGCGCATCGATGGCGCATCGAT 2394
Db 2335 CGCAAGTGTCTTCTGGAGCGCATCGATGGCGCATCGATGGCGCATCGATGGCGCATCGAT 2394
QY 2395 GACCTGTACGTGGCGCGCGCGCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2454
Db 2395 GACCTGTACGTGGCGCGCGCGCTAGGATCGATTAAAGCTTCCCGGGCTAGCAC 2454
QY 2455 GGTGAATTC 2463
Db 2455 GGTGAATTC 2469
```

RESULT 3

ABL39961

ID ABL39961 standard; DNA; 2457 BP.

XX

XX

XX

XX

DT 15-MAY-2002 (first entry)

XX

DE

XX

KW

KW

KW

XX

OS

OS

PN

XX

PD

17-JAN-2002.

Synthetic construct PR975YMMW SEQ ID NO:32.

Human immunodeficiency virus type C; antigenic HIV type C protein;
immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; env; nef;
immunostimulant; gene therapy; gene; ds.Human immunodeficiency virus type C.
Synthetic.

WO200204493-A2.

17-JAN-2002.

XX 05-JUL-2001; 2001WO-US21241.
XX PF
XX PR
XX XX
XX (CHIR) CHIRON CORP.
XX (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
XX in applications including DNA immunization or generation of packaging
XX cell lines, particularly in gene therapy -
XX
XX Claim 1; Fig 10; 233pp; English.
XX
XX The present invention describes expression cassettes comprising a
XX polynucleotide sequence encoding a polypeptide comprising immunogenic
XX HIV type C polypeptides. The expression cassettes comprise any of the
XX HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
XX Nef (1). (1) have immunostimulant activity and can be used in gene
XX therapy. The HIV type C polynucleotides are useful in applications
XX including DNA immunisation, generation of packaging cell lines, and
XX production of HIV type C proteins. The polynucleotides are particularly
XX useful in gene therapy and DNA immunisation applications. ABL39942 to
XX ABL40054 and ABL06204 to ABL06215 represent sequences used in the
XX exemplification of the present invention.
XX
SQ Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 other;

Query Match 98.9%; Score 2436.2; DB 24; Length 2457;
Best Local Similarity 99.6%; Pred. No. 3.2e-257;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

QY 1 GTCAGCGCCACCATGGCGAGGCCATGAGCAGGCCACCGAGCGCAACATCCTGATGCAG 60
DB 1 GTCAGCGCCACCATGGCGAGGCCATGAGCAGGCCACCGAGCGCAACATCCTGATGCAG 60

QY 61 GCGAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGTCTCAACTGGCGCAAGAGGGC 120
DB 61 GCGAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGTCTCAACTGGCGCAAGAGGGC 120

QY 121 CACATCGCCGCCAACTGCGCGGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAG 180
DB 121 CACATCGCCGCCAACTGCGCGGCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAG 180

QY 181 GGCACACAGATGAAGGACTGACCGAGGCGCAGGCCCACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGCACACAGATGAAGGACTGACCGAGGCGCAGGCCCACTTCTTCGCGAGGACCTGGCC 240

QY 241 TTCCCGCCAGGCAAGGCCCGGAGTTCCTCCAGCAGCAGACCGCGCCACAGCCCCACC 300
DB 241 TTCCCGCCAGGCAAGGCCCGGAGTTCCTCCAGCAGCAGACCGCGCCACAGCCCCACC 300

QY 301 AGCGCGAGGTGCAAGTGTGCGGGGCAACCCCGCAGCGAGGCGCGCGCGCGCGCCAG 360
DB 301 AGCGCGAGGTGCAAGTGTGCGGGGCAACCCCGCAGCGAGGCGCGCGCGCGCGCCAG 360

QY 361 GGCACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCCCTTGGTGAAGATCAAGTG 420
DB 361 GGCACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCCCTTGGTGAAGATCAAGTG 420

QY 421 GCGCGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCCGACGACCGCTGCTGGAGGAG 480
DB 421 GCGCGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCCGACGACCGCTGCTGGAGGAG 480

QY 481 ATGAGCTGCGCGGCAAGTGGAGCCCAAGATGATCGCGGCGCATCGCGGCTTTCATCAAG 540
DB 481 ATGAGCTGCGCGGCAAGTGGAGCCCAAGATGATCGCGGCGCATCGCGGCTTTCATCAAG 540

QY 541 GTGCGCCAGTACGACCAAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600

DB 541 GTGCGCCAGTACGACCAAGATCTCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCGTG 600
QY 601 CTGATCGGCCCGCCACCCCGTGAACATCATCGGCCGCAACATCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCCCGCCACCCCGTGAACATCATCGGCCGCAACATCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGGCATG 720
DB 661 ACCCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGGCATG 720
QY 721 GAGCGCCCAAGGTGAAGCAGTGGCCCTGACCCGAGGAGAGATCAAGSCCTTGACCCGCC 780
DB 721 GAGCGCCCAAGGTGAAGCAGTGGCCCTGACCCGAGGAGAGATCAAGSCCTTGACCCGCC 780
QY 781 ATCTGCGAGGAGATGGAGAAGAGGAGGCAAGATCACCAAGATCGGCCCGCAGAACCCCTAC 840
DB 781 ATCTGCGAGGAGATGGAGAAGAGGAGGCAAGATCACCAAGATCGGCCCGCAGAACCCCTAC 840
QY 841 AACACCCCGCTGTTCCGCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGTGGAC 900
DB 841 AACACCCCGCTGTTCCGCCATCAAGAAAGAGACACCAAGTGGCGCAAGCTGTGGAC 900
QY 901 TTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCCAC 960
DB 901 TTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGATCCCCAC 960
QY 961 CCGCGCGCTGGAAGAAGAAAGAGCGTGACCGTGTGGAGCTGGGCGAGCGCTTACTTC 1020
DB 961 CCGCGCGCTGGAAGAAGAAAGAGCGTGACCGTGTGGAGCTGGGCGAGCGCTTACTTC 1020
QY 1021 AGCTGTCCCTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAAC 1080
DB 1021 AGCTGTCCCTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAAC 1080
QY 1081 AACGAGACCCCGCGCATCCGCTACCAAGTACAACGTGCTGCCCGAGGCTGGAAGGCGAGC 1140
DB 1081 AACGAGACCCCGCGCATCCGCTACCAAGTACAACGTGCTGCCCGAGGCTGGAAGGCGAGC 1140
QY 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACCC 1200
DB 1141 CCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCGCTTCCGCGCGCGCAACCC 1200
QY 1201 GAGATCGTGATCTACCAAGCGCCCGCTGACGTGGGACGCGACTGGAGATCGGCCAGCAGC 1260
DB 1201 GAGATCGTGATCTACCAAGCGCCCGCTGACGTGGGACGCGACTGGAGATCGGCCAGCAGC 1260
QY 1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCCACACCCCGCAG 1320
DB 1261 CGCGCCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCCACACCCCGCAG 1320
QY 1321 AAGAAGCACAGAGGAGGCCCGCTTCTGTGGATGGGCTAGGAGCTGCACCCCGCAAG 1380
DB 1321 AAGAAGCACAGAGGAGGCCCGCTTCTGTGGATGGGCTAGGAGCTGCACCCCGCAAG 1380
QY 1381 TGGACCGTGCAGCGCCATCGAGCTGCCGAGAGAGAGCTGACCGTGAACGACATCCAG 1440
DB 1381 TGGACCGTGCAGCGCCATCGAGCTGCCGAGAGAGAGCTGACCGTGAACGACATCCAG 1440
QY 1435 AAGTGTGGCAAGCTGAATGGGCCAGCCAGATCTACCCCGGATCAAGGTGCGCCAG 1500
DB 1435 AAGTGTGGCAAGCTGAATGGGCCAGCCAGATCTACCCCGGATCAAGGTGCGCCAG 1500
QY 1501 CTGTGCAAGCTGTGCGCGCGCCAGGCCCTGACCGACATCTGTGCCCTGACCGAGGAG 1560
DB 1501 CTGTGCAAGCTGTGCGCGCGCCAGGCCCTGACCGACATCTGTGCCCTGACCGAGGAG 1560
QY 1561 GCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGGAGCGCTGCAGCGGCTGTAC 1620
DB 1561 GCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGGAGCGCTGCAGCGGCTGTAC 1620
QY 1621 TACGACCCCGCAGAGGAGCTGTGTGGCGCGAGATCCAGAGAGCGGCGCAGCAGTGGACC 1680
DB 1621 TACGACCCCGCAGAGGAGCTGTGTGGCGCGAGATCCAGAGAGCGGCGCAGCAGTGGACC 1680

Db	1615	TACGACCCGACGAAGGACCTTGGTGGCGGAGATCCAGAACGACGAGCGCCACGACCAAGTGGAC	1674
Qy	1681	TACCAGATCTTACCCAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCAAGATGCGC	1740
Db	1675	TACCAGATCTACCAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCAAGATGCGC	1734
Qy	1741	ACCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCCATGGAG	1800
Db	1735	ACCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAAAGATCGCCATGGAG	1794
Qy	1801	AGCATCTGATCTGGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAAGGAGACCTG	1860
Db	1795	AGCATCTGATCTGGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAAGGAGACCTG	1854
Qy	1861	GAGACCTGGTGGACCGACTACTTGGCAGGCCACCTGATCCCGAGTGGGAGTTCTGTGAAC	1920
Db	1855	GAGACCTGGTGGACCGACTACTTGGCAGGCCACCTGATCCCGAGTGGGAGTTCTGTGAAC	1914
Qy	1921	ACCCCCCCCTGGTCAAGCTGTGGTTACCACTTGGAGAGGAGGCCCATCATCGGCGCCGAG	1980
Db	1915	ACCCCCCCCTGGTCAAGCTGTGGTTACCACTTGGAGAGGAGGCCCATCATCGGCGCCGAG	1974
Qy	1981	ACCTTCTTACGTGGACGGCGCCCAACCGCAGACCAAGATCGGCAAGGCGCGCTACGTG	2040
Db	1975	ACCTTCTTACGTGGACGGCGCCCAACCGCAGACCAAGATCGGCAAGGCGCGCTACGTG	2034
Qy	2041	ACCGACCGGGCGGCAGAAATCTGTGAGCTTGACGAGACCAACCAAGAACGACCGAG	2100
Db	2035	ACCGACCGGGCGGCAGAAATCTGTGAGCTTGACGAGACCAACCAAGAACGACCGAG	2094
Qy	2101	CTGAGGGCATCTACGCTGGCCCTGCAGGACGCGGACGAGGTGAACATCTGTGACCGAC	2160
Db	2095	CTGAGGGCATCTACGCTGGCCCTGCAGGACGCGGACGAGGTGAACATCTGTGACCGAC	2154
Qy	2161	AGCCAGTACGCCCTGGGCATCATCAGGCCCCGACCAAGACGAGCGAGCTGGTG	2220
Db	2155	AGCCAGTACGCCCTGGGCATCATCAGGCCCCGACCAAGACGAGCGAGCTGGTG	2214
Qy	2221	AACCAATCATCGACGACTGATCAAGAGGAGAGGTTACTCTGAGCTGGTGCCCGCC	2280
Db	2215	AACCAATCATCGACGACTGATCAAGAGGAGAGGTTACTCTGAGCTGGTGCCCGCC	2274
Qy	2281	CACRAGGCGATCGCGGCAACGAGCAGATCGACAAGCTGGTGAAGCAAGGCGATCCGCAAG	2340
Db	2275	CACRAGGCGATCGCGGCAACGAGCAGATCGACAAGCTGGTGAAGCAAGGCGATCCGCAAG	2334
Qy	2341	GTGCTGTTCTGGACGGCATCGATGGCGGCATCTGATCTACAGTACATGGACGACCTG	2400
Db	2335	GTGCTGTTCTGGACGGCATCGATGGCGGCATCTGATCTACAGTACATGGACGACCTG	2394
Qy	2401	TACGTGGCAGCGCGGCCCTAGGATTCGATTTAAAGCTTCCCGGGCTAGCACCGGTGA	2460
Db	2395	TACGTGGCAGCGCGGCCCTAGGATTCGATTTAAAGCTTCCCGGGCTAGCACCGGTGA	2454
Qy	2461	TTC 2463	
Db	2455	TTC 2457	

RESULTS

RESULI 4
AAA70479

AAA70479
ID AAA70479 standard: DNA: 2306 BP.

XX
XX
DT
FHM / 04 / 5AAA
AC AAA70479;XX
XX

DT 28-NOV-2000 (first entry)

XX
XX

DE HIV FS(-)_ProtMod_RTopt_YM coding sequence.

XX

KW HIV-1; AIDS; Gag-reverse transcriptase; vac

XX

OS Human immunodeficiency virus type 1.

OS Synthetic.

[illegible]

Db 481 ACCTGCTGACCCAGATCGGCTGCACCCCTGAACCTTCCCATCAGCCCCATCGAGACGGTGC 540
QY 698 CCGTGAAGCTGAAGCCCGGATGACGCGCCCAAGGTGAACACTGCGCCCTGACCGAGG 757
Db 541 CCGTGAAGCTGAAGCCCGGATGACGCGCCCAAGGTGAACACTGCGCCCTGACCGAGG 600
QY 758 AGAAGATCAAGGCCCTGACCCGCTGCGAGGAGATGGAAGAGGAGGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTGCTGGAGATCTGCACCGAGATGGAAGAGGAGGCAAGATCACCA 660
QY 818 AGATCGGCCCGGAGAACCCCTACAAACACCCCGCTGTTGCGCATCAAGAAGAGACAGCA 877
Db 661 AGATCGGCCCGGAGAACCCCTACAAACACCCCGCTGTTGCGCATCAAGAAGAGACAGCA 720
QY 878 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAAGCGACCCAGGACTTCCTGGG 937
Db 721 CCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAAGCGACCCAGGACTTCCTGGG 780
QY 938 AGGTGACAGCTGGGCAATCCCGACCCCGCGCGCTTGAAGAAGAAGAGAGCGTGAACCTGC 997
Db 781 AGGTGACAGCTGGGCAATCCCGACCCCGCGCGCTTGAAGAAGAAGAGAGCGTGAACCTGC 840
QY 998 TGGAGCTGGGCGAGCGCTACTTACAGCTGCCCTGGAGGAGACTTCGCGAAGTACACCG 1057
Db 841 TGGAGCTGGGCGAGCGCTACTTACAGCTGCCCTGGAGGAGACTTCGCGAAGTACACCG 900
QY 1058 CTTTACCATCCCCAGCATCAACAACGAGACCCCGCGCATCCGCTACCAAGTACAACTGC 1117
Db 901 CTTTACCATCCCCAGCATCAACAACGAGACCCCGCGCATCCGCTACCAAGTACAACTGC 960
QY 1118 TGCCCCAGGGCTGAAGGGGAGGCCCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
Db 961 TGCCCCAGGGCTGAAGGGGAGGCCCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020
QY 1178 AGCCCTTCCGCGCGCAACCCAGATCTGTGATCTACCAAGGCCCGCTGTACGTGGCA 1237
Db 1021 AGCCCTTCCGCGCGCAACCCAGATCTGTGATCTACCAAGGCCCGCTGTACGTGGCA 1080
QY 1238 GCGACCTGGAGATCGGCGAGCAGCAGCGCCCAAGATCGAGGAGCTCGCAAGCACCTGCTGC 1297
Db 1081 GCGACCTGGAGATCGGCGAGCAGCAGCGCCCAAGATCGAGGAGCTCGCAAGCACCTGCTGC 1140
QY 1298 GCTGGGGCTTCAACACCCCGGACAAAGACCAAGACCAAGAGGAGCCCGCTTCTGTGATGG 1357
Db 1141 GCTGGGGCTTCAACACCCCGGACAAAGACCAAGACCAAGAGGAGCCCGCTTCTGTGATGG 1200
QY 1358 GCTACGAGCTGCACCCCGCAAGTGAGCGTGGACGCCATCGAGCTGCCCGAGAGAGCA 1417
Db 1201 GCTACGAGCTGCACCCCGCAAGTGAGCGTGGACGCCATCGAGCTGCCCGAGAGAGCA 1260
QY 1418 GCTGGACCGTGAACGACATCCAGAGCTGGTGGCAAGCTGAAGTGGCGCCAGCCAGATCT 1477
Db 1261 GCTGGACCGTGAACGACATCCAGAGCTGGTGGCAAGCTGAAGTGGCGCCAGCCAGATCT 1320
QY 1478 ACCCGGCATCAAGTGGCGGAGCTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCG 1537
Db 1321 ACCCGGCATCAAGTGGCGGAGCTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCG 1380
QY 1538 ACATCGTCCCTGACCGAGGAGCGGAGCTGGAGCTGGCGCGAGAACCGGAGATCTGTC 1597
Db 1381 AGGTGATCCCCCTGACCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGTA 1440
QY 1598 GCGAGCCCGTGCACGGCGTGTACTACACCCCGCAAGAGGAGCTGGTGGCGGAGATCCAGA 1657
Db 1441 AGGAGCCCGTGCAGGAGTGTACTACACCCCGCAAGAGGAGCTGGTGGCGGAGATCCAGA 1500
QY 1658 AGCAGGCCACGACAGCTGACCTTACAGATCTACCAAGAGGCCCTTCAAGAACTGAAGA 1717
Db 1501 AGCAGGCCACGACAGCTGACCTTACAGATCTACCAAGAGGCCCTTCAAGAACTGAAGA 1560
QY 1718 CCGCAAGTACGCAAGATCGGACCGCCCGCACCAACCAAGAGCTGAAGCAGCTGACCGAGG 1777
Db 1561 CCGCAAGTACGCGGATCGCGGGCGCCCGCACCAACCAAGAGCTGAAGCAGCTGACCGAGG 1620

QY 1778 CCGTCCAGAAAGATCCCATGGAGAGCATCGTATCTGGGGCAAGACCCCAAGTTCCGCC 1837
Db 1621 CCGTCCAGAAAGATCGACACCGAGAGCATCGTATCTGGGGCAAGATCCCAAGTTCAAGC 1680
QY 1838 TGCCCATCCAGAAGAGAGACCTGGTGGAGACCTGGTGGACCGACTACTTGGCAGGCGCACTGGA 1897
Db 1681 TGCCCATCCAGAAGAGAGACCTGGTGGAGGCTGGTGGATGGATGCTGGCAGGCGCACTGGA 1740
QY 1898 TCCCGAGTGGAGTCTGTGAACACCCCGCTGGTGAAGCTGGTACCAAGTGGGAGA 1957
Db 1741 TCCCGAGTGGAGTCTGTGAACACCCCGCTGGTGAAGCTGGTACCAAGTGGGAGA 1800
QY 1958 AGGAGCCCATCATCGGCGCGGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCA 2017
Db 1801 AGGAGCCCATCATCGGCGCGGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCA 1860
QY 2018 AGATCGGCAAGCGGCTTACGTGACCGACCGGCGCGGAGATCTGTGAGCTTGACCG 2077
Db 1861 AGCTGGGCAAGCGGCTTACGTGACCGACCGGCGCGGAGAGTGGTGGATCTGCGCG 1920
QY 2078 AGACCAACCAAGAGACCGAGCTGACGAGCCATCCAGCTGGCCCTGAGGACAGCGCA 2137
Db 1921 ACACCAACCAAGAGACCGAGCTGACGAGCCATCCAGCTGGCCCTGAGGACAGCGCGC 1980
QY 2138 GCGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCTCCAGGCCAGCCCG 2197
Db 1981 TGGAGGTGAACATCGTGACCGACAGCCAGTACGCCCTGGGCATCTCCAGGCCAGCCCG 2040
QY 2198 ACAAGAGCGAGGAGCTGTGTGAACCAAGATCATCGACAGCTGATCAAGAGAGAGAGG 2257
Db 2041 ACAAGAGCGAGGAGCTGTGTGAACCAAGATCATCGACAGCTGATCAAGAGAGAGAGG 2100
QY 2258 TGTACCTGAGCTGGTGGCCCGCCCAAGGAGCATCGGGCGCAACGAGCAGATPCGACAGC 2317
Db 2101 TGTACCTGAGCTGGTGGCCCGCCCAAGGAGCATCGGGCGCAACGAGCAGTGGCAAGC 2160
QY 2318 TGGTGAAGCAAGGAGCTCGCAAGTGTCTTCCCTGGACGGATCGATGCGCGCATCTGTA 2377
Db 2161 TGGTGAAGCGCGCATCCGCAAGTGTCTTCCCTGGACGGATCGATGCGCGCATCTGTA 2220
QY 2378 TCTACCAAGTACATGACGACCTGTACCTGGCAGCGCGGCGCTAGGATCGATTAAGAGC 2437
Db 2221 TCTACCAAGTACATGACGACCTGTACCTGGCAGCGCGGCGCTAGGATCGATTAAGAGC 2280
QY 2438 TTCCCGGGGCTAGCACCGGTGAATTC 2463
Db 2281 TTCCCGGGGCTAGCACCGGTGAATTC 2306

RESULT 5

ABK91616

ID - ABK91616 standard; DNA; 9166 BP.

XX ABK91616;

AC AC

XX XX

DT 14-AUG-2002 (first entry)

XX

DE Modified HIV protein-encoding plasmid DNA #168.

XX

KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;

Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

OS

XX WO200232943-A2.

PN

XX

XX

PD 25-APR-2002.

XX

PF 14-AUG-2001; 2001WO-US25721.

XX

PR 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADABARTI B K.
XX
XX Nabel GJ, Huang Y;
XX
DR WPI; 2002-452382/48.
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
PS Disclosure; Page 766-769; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9166 BP; 2227 A; 2662 C; 2513 G; 1764 T; 0 other;

Query Match 82.3%; Score 2027.6; DB 24; Length 9166;
Best Local Similarity 91.7%; Pred. No. 4e-246;
Matches 2194; Conservative 0; Mismatches 179; Indels 19; Gaps 4;

QY 14 TGCCGAGGCGCATGAGCCAGGCCACCGCCCAACATCTCTGATGACGCGCAGCAACTTCA 73
DB 2961 TGCCGAGGCGCATGAGCAAGGTGAACAACCAACATCATGATGACGCGCAGCAACTGCA 3020

QY 74 AGGCCCCCAAGCCCATCATCAAGTGTCTCAACTGCGCGAGGAGGCCACATCGCCCGCA 133
DB 3021 AGGCCCCCAAGCCCATCATCAAGTGTCTCAACTGCGCGAGGAGGCCACATCGCCCGCA 3080

QY 134 ACTGCCGCGCCCCCGCCAGAAAGGGCTGCTGGAAGTTCGCGCAAGGAGGCCACAGATGA 193
DB 3081 ACTGCCGCGCCCCCGCCAGAAAGGGCTGCTGGAAGTTCGCGCAAGGAGGCCACAGATGA 3140

QY 194 AGGACTGACCCAGCGCAGGCCAATCTTCTCCGCGAGGACCTGCGCTTCCCCCAGGGCA 253
DB 3141 AGGACTGACCCAGCGCAGGCCAATCTTCTCCGCGAGGACCTGCGCTTCCCCCAGGGCA 3199

QY 254 AGGCCCGCGAGTTCCCGCAGGAGCAGAACCGCCCAACAGCCGCCACAGCGCGAGCTGC 313
DB 3200 AGGCCAGGGAATTTCTTCAGACGACACACAGACAGCCACCCACAGAGAGAGCTTC 3259

QY 314 AGGTGCGCGG-----CGAACACCCCGCAGGAGGCCGCGCGCAGGGCA--- 364
DB 3260 AGGTTTGGGAAGAGAGAACAACTCCTCTCAGAAAGCAGGAGCGGATAGACAAGGAAGT 3319

QY 365 ---CCCTGAACATTCCTCCAGATCACCTGTGGCAGGCGCCCTTGGTGAGCATCAAGTGG 421
DB 3320 TATCTCTTATAGTTCCTCAGATCACTCTTTGGCAGGCGACCCCTCGTCAATAAAGATAG 3379

QY 422 GCGGCCAGATCAAGAGGCGCTCTTGACACCGCGCGCAGACACCGTCTGAGGAGGA 481
DB 3380 GGGGCCAGCTAAGAGGCGCTCTTACACACCGCGCGCAGACACCGTCTGAGGAGGA 3439

QY 482 TGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGGCGCATCGGCGCTTCAATCAAG 541
DB 3440 TGAACCTGCCCGCGCTGGAAGCCCAAGATGATCGGCGCATCGGCGCTTCAATCAAG 3499

QY 542 TGGCCAGATGACGACAGATCTCTGATGAGATCTGCGGCAAGAGGCGCATCGGCAAGTGC 601
DB 5480 TGTACTACGACCCAGCAAGGACCTGTATCTGCGGAGATCCAGAGGCGGCGGCGGCGAGT 4639

DB 3500 TGGGCCAGTACGACCAGATCTGATCGAGATCTGCGGCCACAAGGCCATCGGCAACCGTGC 3559

QY 602 TGATCGGCGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGTGCA 661
DB 3560 TGGTGGGCGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGATCGGTGCA 3619

QY 662 CCCTGAACCTTCCCATCAGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCCCGGATGG 721
DB 3620 CCCTGAACCTTCCCATCAGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCCCGGATGG 3679

QY 722 ACGGCCCCAAGGTGAAGAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACGCCCA 781
DB 3680 ACGGCCCCAAGGTGAAGAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGCTGAGA 3739

QY 782 TCTGCGAGGAGATGGAAGAGGAGGCAAGATCACCAAGATCGGCCGCGAGAACCCCTACA 841
DB 3740 TCTGCGAGGAGATGGAAGAGGAGGCAAGATCACCAAGATCGGCCGCGAGAACCCCTACA 3799

QY 842 ACACCCCGCTGTTGCGCATCAAGAAAGAAAGGACAGCACCAAGTGGCGCAAGCTGGTGA 901
DB 3800 ACACCCCGCTGTTGCGCATCAAGAAAGAAAGGACAGCACCAAGTGGCGCAAGCTGGTGA 3859

QY 902 TCCGCGAGCTGAACAGCGCACCCAGACTTCTGGAGGTGACAGCTGGGATCCCCCACC 961
DB 3860 TCCGCGAGCTGAACAGCGCACCCAGACTTCTGGAGGTGACAGCTGGGATCCCCCACC 3919

QY 962 CGCCCGGCTGAAGAAAGAAAGAGGAGTGCCTGCTGGAGCTGGGCGAGCCCTACTTCA 1021
DB 3920 CGCCCGGCTGAAGAAAGAAAGAGGAGTGCCTGCTGGAGCTGGGCGAGCCCTACTTCA 3979

QY 1022 GCGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGCAGCATCAACA 1081
DB 3980 GCGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGCAGCATCAACA 4039

QY 1082 ACAGAGCCCCCGGATCCGCTACAGTACAAAGTGTGCTGCCCCAGGCTGGAAGGGCAGCC 1141
DB 4040 ACAGAGCCCCCGGATCCGCTACAGTACAAAGTGTGCTGCCCCAGGCTGGAAGGGCAGCC 4099

QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACCCG 1201
DB 4100 CCAGCATCTTCCAGTGCAGTACCCNAGATCTTGGAGCCCTTCCGCAAGCAGNACCCG 4159

QY 1202 AGATCGTGAATCTACCA-----GGCCCCCTGTACGTGGCAGCGAGCTGAGAGTGGCC 1255
DB 4160 ACATCGTGAATCTACCAATACATGGACCACTGTACGTGGCAGCGAGCTGAGAGTGGCC 4219

QY 1256 AGCACCGCCCAAGATCGAGGAGCTGGCAAGCAGCTGCTGCGCTGGGCTTCCACCAACC 1315
DB 4220 AGCACCGCCCAAGATCGAGGAGCTGGCCAGCAGCTGCTGCGCTGGGCTTCCACCAACC 4279

QY 1316 CCAGCAAGAGACACCAAGAGGAGCCCTTCTCTGTGGATGGGCTAGAGCTGACACCCG 1375
DB 4280 CCAGCAAGAGACACCAAGAGGAGCCCTTCTCTGTGGATGGGCTAGAGCTGACACCCG 4339

QY 1376 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAAGAGAGCTGACCGTGAACGACA 1435
DB 4340 ACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAAGAGAGCTGACCGTGAACGACA 4399

QY 1436 TCCAGAGCTGCTGGGCAAGCTGAACCTGGCCGCGCAGCCAGATCTACCCCGGATCAAGTGC 1495
DB 4400 TCCAGAGCTGCTGGGCAAGCTGAACCTGGCCGCGCAGCCAGATCTACCCCGGATCAAGTGC 4459

QY 1496 GCCAGCTGTCAAGCTGCTGCGGCGGCCCAAGGCCCTGACCCGACATGCTGCCCTGACCG 1555
DB 4460 GCCAGCTGTCAAGCTGCTGCGGCGGCCCAAGGCCCTGACCCGAGGTGCTGCCCTGACCG 4519

QY 1556 AGAGGCGGAGCTGGAGCTGGCGAGAAACCGCGAGATCTCTGCGGAGCCCGCTGACCGGCG 1615
DB 4520 AGAGGCGGAGCTGGAGCTGGCGAGAAACCGCGAGATCTCTGAGAGGAGCCCGCTGACCGG 4579

QY 1616 TGTACTACGACCCCAAGGAGCTGCTGGCCGAGATCCAGAGAGGCGCCACGACAGT 1675
DB 4580 TGTACTACGACCCCAAGGAGCTGCTGGCCGAGATCCAGAGAGGCGCCACGACAGT 4639

PT Gag-containing polypeptide useful for vaccinating against HIV
PT infections and acquired immunodeficiency syndrome (AIDS) -
XX
XX
XX Claim 7; Fig 74; 391pp; English.

CC The present sequence is the coding sequence of a synthetic HIV
CC Gag-reverse transcriptase expression cassette, FS(-).ProtMod.Rtopt_YMMW.
CC The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the Gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC Gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS).

XX
SQ Sequence 2300 BP; 528 A; 754 C; 725 G; 293 T; 0 other;

Query Match 82.0%; Score 2019.2; DB 21; Length 2300;
Best Local Similarity 93.2%; Pred. No. 5.4e-245;
Matches 2150; Conservative 0; Mismatches 138; Indels 18; Gaps 3;

Qy 170 GGGCAAGGAGGCGCCACAGATGAGGAGTGCACCGAGCGCCAGCCAACTTCTTCGCG 229
Db 1 GGGCGCGAGGACACCAATGAAGATTGCATGAGAGAGAGGCTAATTTCTTCGCGG 60
Qy 230 AGGACTGGCTTCCCGCAGGCAAGGCCGCGAGTTCCCGAGAGAGCAACCGCGCCA 289
Db 61 AGGACTGGCTTCCCGCAGGCAAGGCCGCGAGTTCCAGCAGGAGAGACCCGCGCCA 120
Qy 290 ACAGCCCGCAGCGCGAGCTGCGAGTGGCGGGG-----ACAAACCCCGCAGCGAGG 343
Db 121 ACAGCCCGCAGCGCGAGCTGCGAGTGGCGGGGCGGAGAACACAGCTGAGCGAGG 180
Qy 344 CGGCGCGAGCGCCAGGCGACCCCTG-----AACTTCCCGCAGATCACCTGTGGCAGC 397
Db 181 CGGCGCGAGCGCCAGGCGACCTGAGCTTCACTTCCCGCAGATCACCTGTGGCAGC 240
Qy 398 GCCCGCTGGTGAGCATCAAGTGGCGGCGCGAGATCAAGGAGGCCCTGTGGACACCGCG 457
Db 241 GCCCGCTGGTGAGCATCAAGTGGCGGCGCGAGTCAAGGAGCGCTGTGGACACCGCG 300
Qy 458 CCGAGCACCGCTGTGGAGGAGATGACCTGGCGCGAGTACGACAGATCCTGATCGAGATCTGG 517
Db 301 CCGAGCACCGCTGTGGAGGAGATGAACTTCCCGGCAAGTGGAGGCCCAAGATGATCG 360
Qy 518 CGCGATCGCGGCTTCAAGTGGCGCGAGTACGACAGATCCTGATCGAGATCTGG 577
Db 361 CGGGATCGGGGCTTCAAGTGGCGCGAGTACGACAGATCCTGATCGAGATCTGG 420
Qy 578 GCAAGAGGCCATCGGCAACCGCTGTGATCGGCGCGACCCCGGTGAACATCATCGGCGCA 637
Db 421 GCCAAGAGGCCATCGGCAACCGCTGTGATCGGCGCGACCCCGGTGAACATCATCGGCGCA 480
Qy 638 ACATGCTACCCAGCTGGGCTGCACTTCACTTCCCGATAGCCCGATCGAGACCGTGC 697
Db 481 ACCTGCTACCCAGATCGGCTGCACTTCACTTCCCGATAGCCCGATCGAGACCGTGC 540
Qy 698 CCGTGAAGCTGAAGCCCGGCTGAGCGCGCGAGTGAAGTGGCGCGCTGACCGAGG 757
Db 541 CCGTGAAGCTGAAGCCCGGCTGAGCGCGCGAGTGAAGTGGCGCGCTGACCGAGG 600
Qy 758 AGAGATCAAGCCCTGACCGCGATCTCGAGGAGATGGAGAGGAGGCGCAAGATCAACA 817
Db 601 AGAGATCAAGCCCTGAGCGGATCTGACCGGAGATGGAGAGGAGGCGCAAGATCAACA 660
Qy 818 AGATGGCGCGAGACCCCTACACACCCCGCTGTCGCCATCAAGAGAGAGACACA 877
Db 661 AGATGGCGCGAGACCCCTACACACCCCGCTGTCGCCATCAAGAGAGAGACACA 720
Qy 878 CCAAGTGGCGCAAGCTGGTGGACTTCCCGAGCTGAACAAGCGACCCAGGACTTCTGGG 937
Db 721 CCAAGTGGCGCAAGCTGGTGGACTTCCCGAGCTGAACAAGCGACCCAGGACTTCTGGG 780

Qy 938 AGGTGAGCTGGGCTATCCCGCCCGCGCTGAAGAAGAGAGAGCGTGACCGTGC 997
Db 781 AGGTGAGCTGGGCTATCCCGCCCGCGCTGAAGAAGAGAGAGCGTGACCGTGC 840
Qy 998 TGGAGCTGGGCGAGCGCTACTTACAGCTGCCCTGACGAGGACTTCCGCAAGTACACCG 1057
Db 841 TGGAGCTGGGCGAGCGCTACTTACAGCTGCCCTGACGAGGACTTCCGCAAGTACACCG 900
Qy 1058 CTTTACCATTCCCGAGCATCAACAGAGACCCCGCGCATCCGCTACCAAGTACAGCTGC 1117
Db 901 CTTTACCATTCCCGAGCATCAACAGAGACCCCGCGCATCCGCTACCAAGTACAGCTGC 960
Qy 1118 TGCCCGAGGCTGGAAGGCGAGCCCGCAGCATCTTCAGAGCAGCATGACCAAGTCTTG 1177
Db 961 TGCCCGAGGCTGGAAGGCGAGCCCGCAGCATCTTCAGAGCAGCATGACCAAGTCTTG 1020
Qy 1178 AGCCCTTCGCGCGCCGCAACCCCGAGATCGTATCTACAGGCGCCCTGTAGTGGGCA 1237
Db 1021 AGCCCTTCGCGCGAGCAACCCCGCATCTATCTACAGGCGCCCTGTAGTGGGCA 1080
Qy 1238 GCGACTGGAGATCGCGCAGCAGCCGCGCAAGATCGAGGCTGCGCAAGCAGCTCTGC 1297
Db 1081 GCGACTGGAGATCGCGCAGCAGCCGCGCAAGATCGAGGCTGCGCGCAGCAGCTCTGC 1140
Qy 1298 GCTGGGCTTACACCCCGCAGACAGAGCAGCAGAGGAGCCCGCTTCTCTGTGATGG 1357
Db 1141 GCTGGGCTTACACCCCGCAGACAGAGCAGCAGAGGAGCCCGCTTCTCTGCCAT-- 1198
Qy 1358 GCTAGAGCTGACCCCGCAGAAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGGA 1417
Db 1199 ---CGAGTGCACCCCGCAGAAAGTGGACCGTGCAGCCCATCTGCTGCCGAGAGGACA 1254
Qy 1418 GCTGACCGTGAACACATCCAGAGCTGGTGGCAAGCTGACTGGGCGCAGCCAGATCT 1477
Db 1255 GCTGACCGTGAACACATCCAGAGCTGGTGGCAAGCTGACTGGGCGCAGCCAGATCT 1314
Qy 1478 ACCCGGCTCAAGTGGCGCGAGCTGTCAAGCTGTGCGCGCGCGCAAGGCGCTGACCG 1537
Db 1315 AGCGCGGCTCAAGTGAAGTGAAGTGTGCAAGCTGCTGCGCGCAGCAGGCGCTGACCG 1374
Qy 1538 ACATGCTGCCCTGACCGAGGAGCGGAGCTGGAGTGGCGGAGAACCGCGAGATCTCTGC 1597
Db 1375 AGGTGATCCCGTGCAGGAGGCGGAGCTGGAGTGGCGGAGAACCGCGAGATCTCTGA 1434
Qy 1598 GGGAGCCGTGCACCGCGTGTACTAGACCCCGCAGAGACCTGTGGCGGAGATCCACA 1657
Db 1435 AGGAGCCGTGCACCGAGGTGTACTAGACCCCGCAGAGACCTGTGGCGGAGATCCACA 1494
Qy 1658 AGCAGGCGCAGCAGCTGGAGCTTACAGATCTACAGAGCGCTTCAAGAACCTGAAGA 1717
Db 1495 AGCAGGCGCAGGCGCAGTGGACCTACAGATCTACAGAGCGCTTCAAGAACCTGAAGA 1554
Qy 1718 CCGGCAAGTACGCAAGATGCGCCACCGCGCGCGCGCACCAACGACGTGAAGCAGTGA 1777
Db 1555 CCGGCAAGTACGCGCGCATGCGCGCGCGCGCGCACCAACGACGTGAAGCAGTGA 1614
Qy 1778 CGTGCAGAAAGTCCCATGGAGCATCTGATCTGGGGCAAGACCCCGCAAGTTCGCGC 1837
Db 1615 CGTGCAGAAAGTGAACCGCAGCATCTGATCTGGGGCAAGATCCCGCAAGTTCGCGC 1674
Qy 1838 TGCCCATCCAGAGGAGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCGCAGCTGA 1897
Db 1675 TGCCCATCCAGAGGAGACCTGGGAGGCTGGTGGATGGATGAGTACTGGCAGGCGCAGCTGA 1734
Qy 1898 TCCCGAGTGGAGTTCGTGAACACCCCGCTGGTGAAGCTGGTACCAAGTGGAGGA 1957
Db 1735 TCCCGAGTGGAGTTCGTGAACACCCCGCTGGTGAAGCTGGTACCAAGTGGAGGA 1794
Qy 1958 AGGAGCCCATCATCGCGCGCAGACCTTCTAGTGGAGCGCGCGCGCAACCCCGGAGACA 2017
Db 1795 AGGAGCCCATCATCGCGCGCGCAGACCTTCTAGTGGAGCGCGCGCGCAACCCCGGAGACA 1854
Qy 2018 AGATCGCAAGGCGCGCTACGTGACCGAGCGCGCGCGAGAGATCGTGAAGCTGACCG 2077

Db 1855 AGCTGGCAAGCGCGCTACGTGACCGACCGCGCGCGAGAGGTGGTGAAGCATCGCG 1914
Qy 2078 AGACCACCAACAGACCGAGCTGCGAGGCCATCCAGCTGGCCCTGCGAGGAGCGGCA 2137
Db 1915 ACACCAACCAACAGACCGAGCTGCGAGGCCATCCAGCTGGCCCTGCGAGGAGCGGCG 1974
Qy 2138 GCGAGGTGAACATCGTGACCGACAGCCAGTACCGCTGGGCATCATCGAGGCCAGCGCG 2197
Db 1975 TGAGGTGAACATCGTGACCGACAGCCAGTACCGCTGGGCATCATCGAGGCCAGCGCG 2034
Qy 2198 ACAAGAGCGAGCGAGCTGGTGAACAGATCATCGAGCGAGCTGATCAAGAAGGAGAGG 2257
Db 2035 ACAAGAGCGAGCGAGCTGGTGAACAGATCATCGAGCGAGCTGATCAAGAAGGAGAGG 2094
Qy 2258 TGACCTGAGCTGGTGGCGCCGCCACAGGCGATCGCGCGAGCGAGCGATCGACAGC 2317
Db 2095 TGTACCTGGCTGGTGGCGCCGCCACAGGCGATCGCGCGAGCGAGCGATCGACAGC 2154
Qy 2318 TGTGAGCAAGGGCATCCGCAAGGTGCTGTTCTGAGCGGCATCGATGGCGGCATCGTA 2377
Db 2155 TGTGAGCGCGGCATCCGCAAGGTGCTGTTCTGAGCGGCATCGATGGCGGCATCGTA 2214
Qy 2378 TCTACCAAGTACATGGACGACCTGTACGTGGGCGAGCGCGCCCTAGGATCGATTAAAGC 2437
Db 2215 TCTACCAAGTACATGGACGACCTGTACGTGGGCGAGCGCGCCCTAGGATCGATTAAAGC 2274
Qy 2438 TTCCCGGGGTAGCACCGGTGAATTC 2463
Db 2275 TTCCCGGGGTAGCACCGGTGAATTC 2300

RESULT 8

ABK91622

ID ABK91622 standard; DNA; 9788 BP.

XX

AC ABK91622;

XX

14-AUG-2002 (first entry)

XX

DE Modified HIV protein-encoding plasmid DNA #174.

XX

KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX

OS Synthetic.

XX

PN WO20232943-A2.

XX

PD 25-APR-2002.

XX

PF 14-AUG-2001; 2001WO-US25721.

XX

PR 14-AUG-2000; 2000US-225097P.

XX

PR 14-NOV-2000; 2000US-252115P.

XX

PR 28-MAR-2001; 2001US-279257P.

XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRABARTI B K.

XX

PI Nabel GJ, Huang Y;

XX

XX WPI; 2002-452382/48.

XX

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and

XX its encoded protein, useful as vaccines for genetic or protein

XX immunisation for acquired immunodeficiency syndrome or human

XX immunodeficiency virus infection

XX

PS Disclosure; Page 785-788; 794pp; English.

XX

CC The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX

SQ Sequence 9788 BP; 2377 A; 2817 C; 2695 G; 1899 T; 0 other;

Query Match 82.0%; Score 2019.2; DB 24; Length 9788;

Best Local Similarity 91.6%; Pred. No. 4.4e-245;

Matches 2190; Conservative 0; Mismatches 178; Indels 24; Gaps 4;

Qy 14 TGGCCGAGGCGCATGAGCGAGGCCACAGCGCCAACTCTGTATCGACGCGCAACTTCA 73

Db 2961 TGGCCGAGGCGCATGAGCGAGGTGAACAACCAACATCATGATCGACGCGCAACTGCA 3020

Qy 74 AGGCCCCCAAGCGCATCATCAAGTCTTCAACTCGGCAAGGAGGCGCACATCGCCGCA 133

Db 3021 AGGCCCCCAAGCGCACATCAAGTCTTCAACTCGGCAAGGAGGCGCACATCGCCGCA 3080

Qy 134 ACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCACAGATGA 193

Db 3081 ACTGCCCGCGCCCGCCCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCGCACAGATGA 3140

Qy 194 AGGACTCACCGAGCGCGAGGCCAACTTCTTCCGCGAGGAGCTTGCCCTTCCCCAGGGCA 253

Db 3141 AGGACTCACCGAGCGAGGGCTAA-----TAGGGAAGATCTGGCTTCCCAAGGGA 3194

Qy 254 AGGCCCGCGAGTCCCGAGCGAGAGAACCGCGCCACACAGCCCGCCAGCGCGAGCTGC 313

Db 3195 AGGCCAGGGAATTTCTTTCAGAGCAGACACAGCCACAGCCCGCCAGAGAGAGCTTC 3254

Qy 314 AGGTGCGCGG-----CGACAACCCCGCGAGCGCGCGCGCGCGAGCGCGCGCA--- 364

Db 3255 AGTTTGGGGAAGAGACAACAATCCCTCTCAGAAAGCAGAGCGCGATAGACAAGAACTG 3314

Qy 365 ---CCCTGAATTCCTCCCGAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGTGG 421

Db 3315 TATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCAGCCCTGCTCAACAATAAGATAG 3374

Qy 422 GCGCCAGATCAAGGAGGCGCTGCTGGACACCGCGCGAGCACCGTGTGAGGAGA 481

Db 3375 GGGGCCAGCTGAGGAGGCGCTTCTAGACACCGCGCGAGCACCGTGTGAGGAGA 3434

Qy 482 TGAGCCTCGCCCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGGGGCTTTCATCAAG 541

Db 3435 TGAACCTGCCCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGGGGCTTTCATCAAG 3494

Qy 542 TGGCCAGTACGACAGATCTGTATCGAGATCTCGGCAAGAGGCCATCGGCGACCGTGC 601

Db 3495 TGGCCAGTACGACAGATCTGTATCGAGATCTCGGCGCAAGAGGCCATCGGCGACCGTGC 3554

Qy 602 TGATCGGCGCCCGCTGNAACATCATCGGCGCAACATCTGACCCAGCTGGGCTGCA 661

Db 3555 TGGTGGGCGCCCGCTGNAACATCATCGGCGCAACATCTGACCCAGCTGGGCTGCA 3614

Qy 662 CCCTGAATTCCTCCCGCATCAGCCCGCATCGAGACCGTGCCTGAAAGTGAAGCCCGCATGG 721

Db 3615 CCCTGAATTCCTCCCGCATCAGCCCGCATCGAGACCGTGCCTGAAAGTGAAGCCCGCATGG 3674

Qy 722 ACGCCCGCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAAGATCAAGGCCCTGACCGCA 781

Db 3675 ACGCCCGCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAAAGATCAAGGCCCTGATGGGAGA 3734

Qy 782 TCTCGAGGAGATGAGAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCTACA 841

Db 3735 TCTCGAGGAGATGAGAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCTACA 3794

QY 842 ACACCCCGTGTTCGCCATCAAGAAAGAGACAGACACCAAGTGGCGCAAGCTGGTGGACT 901
Db ACACCCCGTGTTCGCCATCAAGAAAGAGACAGACACCAAGTGGCGCAAGCTGGTGGACT 3854
QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTCAGCTGGGCATCCCCACC 961
Db TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTCAGCTGGGCATCCCCACC 3914
QY 962 CCGCGCGCTGAAGAAGAAAGAGCGGTGACCGTGTGGAGCTGGGCGACGCCCTACTTCA 1021
Db CCGCGCGCTGAAGAAGAAAGAGCGGTGACCGTGTGGAGCTGGGCGACGCCCTACTTCA 3974
QY 1022 CGTGGCCCTGGAGAGGACTTCCGCAAGTACACGCCCTTCAACATCCCCAGCATCAACA 1081
Db CGTGGCCCTGGAGAGGACTTCCGCAAGTACACGCCCTTCAACATCCCCAGCATCAACA 4034
QY 1082 ACAGAGACCCCGGATCCGCTACCAAGTACACGCTGTCCGCCAGGCTGGAGGCGACGCC 1141
Db ACAGAGACCCCGGATCCGCTACCAAGTACACGCTGTCCGCCAGGCTGGAGGCGACGCC 4094
QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCCTGGAGCCCTTCCGCCCGCCACACCCG 1201
Db CCAGCATCTTCCAGTGCAGCATGACCAAGATCCCTGGAGCCCTTCCGCCCGCCACACCCG 4154
QY 1202 AGATCGTGTACTACA-----GGCCCCCTGTAGTGGGCGAGCACCTGGAGATCGGCC 1255
Db AGATCGTGTACTACA-----GGCCCCCTGTAGTGGGCGAGCACCTGGAGATCGGCC 4214
QY 1256 AGACCGCGCCAAAGTACGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCAACACCC 1315
Db AGACCGCGCCAAAGTACGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCAACACCC 4274
QY 1316 CCGAAGAAGACACCAAGAGAGCCCGCTTCCTGTGTGATGGGCTACGAGCTGCACCCCG 1375
Db CCGAAGAAGACACCAAGAGAGCCCGCTTCCTGTGTGATGGGCTACGAGCTGCACCCCG 4334
QY 1376 ACAAGTGCAGCTGCAGCCCATCGAGCTGCCCGAGAGAGAGAGTGGACCGTGAACGACA 1435
Db ACAAGTGCAGCTGCAGCCCATCGAGCTGCCCGAGAGAGAGAGTGGACCGTGAACGACA 4394
QY 1436 TCCAGAGCTGTGGGCAAGCTGAACCTGGCGACCGCAGATCTACCCGCGCATCAAGTGC 1495
Db TCCAGAGCTGTGGGCAAGCTGAACCTGGCGACCGCAGATCTACCCGCGCATCAAGTGC 4454
QY 1496 GCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTGACGACATGTCGCCCTGACCG 1555
Db GCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTGACGACATGTCGCCCTGACCG 4514
QY 1556 AGGAGCGGAGCTGGAGCTGGCGGAGACCGCGAGATCTCGCGGAGCCCGTGCACGCG 1615
Db AGGAGCGGAGCTGGAGCTGGCGGAGACCGCGAGATCTCGCGGAGCCCGTGCACGCG 4574
QY 1616 TGTACTACGACCCCGACGAGCTGTGGCGGAGATCCAGAAAGCAGGCGCACACAGT 1675
Db TGTACTACGACCCCGACGAGCTGTGGCGGAGATCCAGAAAGCAGGCGCACACAGT 4634
QY 1676 GGACCTACAGATCTACGAGGACCTTCAAGAACCCTGAAGACCGGCAAGTACGCCAAGA 1735
Db GGACCTACAGATCTACGAGGACCTTCAAGAACCCTGAAGACCGGCAAGTACGCCAAGA 4694
QY 1736 TGGCAGCGCCACCAACGAGCTGAAGAGCTGACCGAGGCGCGTGCAGAAATCGCCA 1795
Db TGGCAGCGCCACCAACGAGCTGAAGAGCTGACCGAGGCGCGTGCAGAAATCGCCA 4754
QY 1796 TGGAGAGCATCTGATCTGGGCGAAGACCCCGAGTTCCGCTCGCCATCCAGAGAGA 1855
Db TGGAGAGCATCTGATCTGGGCGAAGACCCCGAGTTCCGCTCGCCATCCAGAGAGA 4814
QY 1856 CCTGGGAGACCTGGTGACCGACTTACCTGGCAGGCGACCTGGATCCCGAGTGGAGTTCG 1915
Db CCTGGGAGGCTGGTGACCGAGTACTGGCAGGCGACCTGGATCCCGAGTGGAGTTCG 4874
QY 1916 TGAACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCCATCATCGCG 1975

Db 4875 TGAACACCCCGCTGGTGAAGCTGTGGTACCAGCTGAGAAGAGCCCATCATCGCG 4934
QY 1976 CCGAGACCTTACGTGGACGCGCGCCCAACCGGAGACCAAGATCGGAAGCGCGGT 2035
Db CCGAGACCTTACGTGGACGCGCGCGCCCAACCGGAGACCAAGATCGGAAGCGCGGT 4994
QY 2036 ACGTGACCGACCGGCGCGGAGAGATCGTGAGCTGACCGAGACCAACCAAGAGA 2095
Db ACGTGACCGACCGGCGCGGAGAGATCGTGAGCTGACCGAGACCAACCAAGAGA 5054
QY 2096 CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCGTA 2155
Db CCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCGTA 5114
QY 2156 CCGACAGCAGTACGCTGGGCGCATCCAGGCGCGCGCGCGAGAGAGAGGAGC 2215
Db CCGACAGCAGTACGCTGGGCGCATCCAGGCGCGCGCGCGAGAGAGAGGAGC 5174
QY 2216 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGC 2275
Db TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGC 5234
QY 2276 CCGCCCAAGAGGATCGGCGGCAACGAGCAGATCGACAAGCTGCTGAGCAAGGCGATCC 2335
Db CCGCCCAAGAGGATCGGCGGCAACGAGCAGTGTGAGCGGCTGCTGAGCGCGGCGATCC 5294
QY 2336 GCAAGGTGCTTCTCTGACGCGCATCGATGGGCGCATGCTGATCTACAGTA 2387
Db GCAAGGTGCTTCTCTGACGCGCATCGATGGGCGCGCGCGAGGAGGAGCGAGAA 5346

RESULT 9

ABK91617

ID ABK91617 standard; DNA; 9169 BP.

XX ABK91617;

XX AC

XX 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #169.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

XX WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

XX 14-NOV-2000; 2000US-252115P.

XX 28-MAR-2001; 2001US-279257P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (CHAD/) CHADRABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
its encoded protein, useful as vaccines for genetic or protein
immunisation for acquired immunodeficiency syndrome or human
immunodeficiency virus infection

XX Disclosure; Page 769-772; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

(human immunodeficiency virus) protein. The nucleic acid molecule and its encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for genetic or protein immunisation to a host, respectively. In particular these are useful for ameliorating the symptoms of acquired immunodeficiency syndrome (AIDS) or HIV infection and generating an antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env, Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent compositions are useful for treating or preventing HIV infections or AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding plasmid DNA sequences of the invention.

Sequence 9169 BP; 2227 A; 2668 C; 2505 G; 1769 T; 0 other;

Query Match	81.4%;	Score	2005.2;	DB	24;	Length	9169;
Best Local Similarity	91.4%;	Pred. No.	2.5e-243;				
Matches 2187; Conservative	0;	Mismatches	183;	Indels	22;	Gaps	5;

Qy	14	TGGCCGAGCCATGAGCCAGGCCACAGGGCCCAACATCCTGATGACGGCAGCAACTTCA	73
Db	2967	TGGCCGAGGCCATGAGCCAGGTCGACAGCAACCAACATCATGATGACGGCGGCAACTTCA	3026
Qy	74	AGGCCCCAAAGCCATCATCAAGTGTCTCAACTGGGCAAGAGGGCCACATGCGCCGCA	133
Db	3027	AGGCCAGAGCCCAT--CAAGTGTCTCACTGGGCAAGAGGGCCACTGGCCGCA	3083
Qy	134	ACTGCGCGCCCCGCCAAGAAAGGGCTGCTGGAAGTCGGGCAAGAGGGCCACACAGATGA	193
Db	3084	ACTGCGCGCCCTGCGCAAGAAAGGGTGCTGGAAGTCGGGCAAGAGGGCCACCAGATGA	3143
Qy	194	AGGACTGCACCGAGCGCCAGGCCAATCTTCCGCGAGGACCTGGCCTTCCCCAGGGCA	253
Db	3144	AGGACTGCACCGAGCGACAGGGTAA--TTTTTTAGGGAAGATCTGGCCTTCCCAACAGGA	3202
Qy	254	AGGCCCGCAGTTTCCCGACGAGCAGAAACCGCGCCAAACAGACCCACCAGCGCGAGCTGC	313
Db	3203	AGGCCAGGGAATTTTCTCAGAGCAGACAGAGCCACAGCCCCACAGAGAGAGCTTC	3262
Qy	314	AGTGTGCGGG-----CGACACCCCGCAGGAGGGCGGGCGAGGGCCAGGGCA---	364
Db	3263	AGGTTTGGGGAAGAGACACAACCTCCCTCTCAGAAGCAGAGCGCGATAGACAGAGAACTG	3322
Qy	365	---CCCTGAACTTCCCGCCAGATCACCTGTGCGACGCGCCCTGTGTAGAGTCAAGGTGG	421
Db	3323	TATCTTTAGCTTCCCTCAGATCACTCTTTGCGACCGACCCCTCGTCAATAAAGATAG	3382
Qy	422	CGCGCCAGATCAAGAGAGGCCCTGCTGGACACCGCGCGCCGACACACCTGTGTGAGGAGA	481
Db	3383	GGGSCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCAGACACCGTGTGTGAGGAGA	3442
Qy	482	TGACCTGCCCCGCAAGTGGAGGCCCAAGATGATCGGCGGCATCGGCGGCTTTCATCAAG	541
Db	3443	TGAACCTTGCCCGCGCGCTGGAGGCCCAAGATGATCGGCGGCATCGGCGGCTTTCATCAAG	3502
Qy	542	TGGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAAAGGCCATCGGACCGTGC	601
Db	3503	TGGCCAGTACGACAGATCTCTGATCGAGATCTGCGCGCCACAGGCCCATCGGACCGTGC	3562
Qy	602	TGATCGGCCCCACCCCGCTGAACATCATCGGCGCCGAACATGCTGACCCAGCTGGGCTGCA	661
Db	3563	TGTTGGGCCCCACCCCGGTGAACATCATCGGCGCGCACTGCTGACCCAGATCGGCTGCA	3622
Qy	662	CCCTGAACCTTCCCATCAGCCCCATCGAGACCGTGCCTGCTGAAGCTGAAGCCCGCATGG	721
Db	3623	CCCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCTGCTGAAGCTGAAGCCCGGCATGG	3682
Qy	722	ACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCGCCA	781
Db	3683	ACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGTCAAGGCCCTGTTGGAGA	3742
Qy	782	TCTGCGAGGAGATGGAGAAGGAGGCGAGATCACCAAGATCGGCCCCCGAGAACCCCTTACA	841
Db	3743	TCTGCACCGAGATGGAGAAGGCGAGATCAGCAAGATGGCCCCGAGAACCCCTTACA	3802

Db 4883 TGAACACCCCCCTGCTGAAGCTGTGTTACCAAGTGGAGAGGAGCCCATCATCGGG 4942
QY 1976 CCGAGACCTTCTAGCTGACGGCGCCGACCAAGATCGGCAAGATCGGCAAGCGCGCT 2035
Db 4943 CCGAGACCTTCTAGCTGACGGCGCCGACCAAGATCGGCAAGATCGGCAAGCGCGCT 5002
QY 2036 ACCTGACCGACCGGGCGCGGAGAGATCGTGAAGCTGACCGAGACCAACCAACGAAGA 2095
Db 5003 ACCTGACCGACCGGGCGCGGAGAGTGTGCTGCTGACCGACCAACCAACGAAGA 5062
QY 2096 CCGAGCTGACGGCATCAGCTGGCCCTGAGGACAGCGGAGCGGAGGATGACATCGTGA 2155
Db 5063 CCGAGCTGACGGCATCAGCTGGCCCTGAGGACAGCGGAGCGGAGGATGACATCGTGA 5122
QY 2156 CCGAGACGACGCTGCGGCTGATCATCCAGGCGCCGACCAAGAGCGAGCGGAGCGGAGC 2215
Db 5123 CCGAGACGACGCTGCGGCTGATCATCCAGGCGCCGACCAAGAGCGGAGCGGAGC 5182
QY 2216 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTGGGTGC 2275
Db 5183 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTGGGTGC 5242
QY 2276 CCGCCCAACAGGCGATCGGCGGCAACGAGAGATCGACAAAGCTGGTGAAGGCGCATCC 2335
Db 5243 CCGCCCAACAGGCGATCGGCGGCAACGAGAGTGGAGCGGCTGGTGAAGGCGCATCC 5302
QY 2336 GCAAGGTGCTCTTCTGACGGCATCGATGCGGCGCATCGTATCATACAGTA 2387
Db 5303 GCAAGGTGCTCTTCTGACGGCATCGATGCGGCGCATCGTATCATACAGTA 5354

RESULT 10
ABK91612
ID ABK91612 standard; DNA; 9194 BP.
XX AC ABK91612;
XX 14-AUG-2002 (first entry)
DE Modified HIV protein-encoding plasmid DNA #164.
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX Synthetic.
XX WO200232943-A2.
XX 25-APR-2002.
XX 14-AUG-2001; 2001WO-US25721.
XX 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
XX 28-MAR-2001; 2001US-279257P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRABARTI B K.
XX Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX Disclosure; Page 753-756; 794pp; English.
XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX

SQ Sequence 9194 BP; 2232 A; 2672 C; 2523 G; 1767 T; 0 other;
Query Match 81.3%; Score 2001.8; DB 24; Length 9194;
Best Local Similarity 91.3%; Pred. No. 6.8e-243;
Matches 2186; Conservative 0; Mismatches 187; Indels 22; Gaps 5;

QY 14 TGGCGAGCGCATGAGCCAGG---CCAGCAGCGCCCAACATCTGATGAGCGCAGCAACT 70
Db 2967 TGGCGAGCGCATGAGCCAGGTTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026
QY 71 TCAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGAGGCGCCACATCGGCC 130
Db 3027 TCCGCAACACAGCGCAAGATCGTGAAGTGTCTCAACTGCGGCAAGAGGCGCCACCGGCC 3086
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGGCAAGAGGCGCCACACCA 190
Db 3087 GCAACTGCGCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGGCAAGAGGCGCCACACCA 3146
QY 191 TGAAGGACTGACCGAGCGCCAGGCAACTTCTTCCGCGAGGACCTGGCTTCCCGCCAGG 250
Db 3147 TGAAGGACTGACCGAGCGCCAGGCTAA-TTTTATGGAAGATCTGGCTTCCACACAG 3205
QY 251 GCAAGGCCCGGAGTTTCCCGAGCAGCAGAACCGCGCCCAACAGCCCCACAGCCCGCAGC 310
Db 3206 GGAAGGCCAGGGAATTTTCTCAGAGCAGACAGACCAACAGCCCGCCACAGAGAGAGC 3265
QY 311 TGCAGGTGCGCGG---CGACAAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCGCGCA 364
Db 3266 TGCAGGTGCGCGGAAAGAGACAACTCCCTCTCAGAAAGAGGAGCGCATGAGCAAGGAA 3325
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTTGTGACGCGCCCTGGTGTGACATCAAGG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCTCTGCTCACAATAAGA 3385
QY 419 TGGCGGCGCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGCAGCAGCACCCTGTGGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGCAGCAGCACCCTGTGGAGG 3445
QY 479 AGATGAGCTGCGCGCAAGTGAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 538
Db 3446 AGATGAGCTGCGCGCGCTGGAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 3505
QY 539 AGGTGCGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGCCCATCGGCACCG 598
Db 3506 AGGTGCGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCGCACAAGCCCATCGCACCG 3565
QY 599 TGCTGATCGGCGCCACCGCGCTGACATCATTCGCGCGCAACATGCTGACCCAGTGGGCT 658
Db 3566 TGCTGATCGGCGCCACCGCGCTGACATCATTCGCGCGCAACATGCTGACCCAGATCGGCT 3625
QY 659 GCACCTGAACTTCCCATCATGAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGCA 718
Db 3626 GCACCTGAACTTCCCATCATGAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGCA 3685
QY 719 TGGAGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGTACCG 778
Db 3686 TGGAGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGTGGTG 3745
QY 779 CCATCTCGGAGGAGATGGAAGAGGCGCAAGATCAACCAAGATCGGCGCGCGAGAACCCCT 838
Db 3746 AGATCTCGGAGGAGATGGAAGAGGCGCAAGATCAACCAAGATCGGCGCGCGAGAACCCCT 3805

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 9194 BP: 2232 A; 2671 C; 2524 G; 1767 T; 0 other;

Query Match 81.2%; Score 2000.2; DB 24; Length 9194;
Best Local Similarity 91.2%; Pred. No. 1.1e-242;
Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;

QY 14 TGGCCGAGGCCATGAGCCAGG---CCACGAGCGCCACATCTCTGATGAGCGCAGCAACT 70
Db TGGCCGAGGCCATGAGCGAGCTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGCCTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
Db TGGCCGAGGCCATGAGCGAGCTGACCAACAGCGCCACATCATGATGAGCGCGCAACT 3026
QY 3027 TCOCACACGAGCGCAAGATCGTGAAGTGCCTTCAACTGCGCAAGGAGGGCCACACCGCCC 3086
Db TCOCACACGAGCGCAAGATCGTGAAGTGCCTTCAACTGCGCAAGGAGGGCCACACCGCCC 3086
QY 131 GCAACTGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
Db GCAACTGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGACGAGCGCGCCAGGCCAATCTTTCGCGGAGGACCTGGCCCTTCCCGCAGG 250
Db TGAAGGACTGACGAGCGCGCCAGGCCAATCTTTCGCGGAGGACCTGGCCCTTCCCGCAGG 3205
QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGCAGAACCGCGCCACAGCGCCACACCGCGCAGC 310
Db GCAAGGCGCGGAGTTCGCCAGCGAGCAGAACCGCGCCACAGCGCCACACCGCGCAGC 310
QY 3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACACAGGCCAACAGCGCCACACAGAGAGC 3265
Db GGAAGGCCAGGGAATTTCTTCAGAGCAGACACAGGCCAACAGCGCCACACAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCGCCAGGCA 364
Db TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGAGCGCGCCAGGCA 364
QY 3266 TTCAAGGTTTGGGGAAGAGAAACAACCTCCCTCTTCAGAAGCAGGAGCGGATAGACAAGAA 3325
Db TTCAAGGTTTGGGGAAGAGAAACAACCTCCCTCTTCAGAAGCAGGAGCGGATAGACAAGAA 3325
QY 365 -----CCCTGAATTCCTCCCGACATCACTCTGGCAGCGCCCTTGGTGAAGATCAAGG 418
Db -----CCCTGAATTCCTCCCGACATCACTCTGGCAGCGCCCTTGGTGAAGATCAAGG 418
QY 3326 CTGTATCTTTAGCTTCCCTTCAGATCACTCTTTGGCAGCAGCCCTCTGCATCAATAAGA 3385
Db CTGTATCTTTAGCTTCCCTTCAGATCACTCTTTGGCAGCAGCCCTCTGCATCAATAAGA 3385
QY 419 TGGGCGGCCAGATCAAGGAGCGCCTGCTGGACACCGCGCGCAGCACACCGTGTGTGAGG 478
Db TGGGCGGCCAGATCAAGGAGCGCCTGCTGGACACCGCGCGCAGCACACCGTGTGTGAGG 478
QY 3386 TAGGGGCCAGCTGAAGGAGGCCCTTCTAGACACCGCGCGGAGCAGACCGTGTGTGAGG 3445
Db TAGGGGCCAGCTGAAGGAGGCCCTTCTAGACACCGCGCGGAGCAGACCGTGTGTGAGG 3445
QY 479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGGCGGCATCGCGGCTTCATCA 538
Db AGATGAGCCTGCCCGGCGCTGGAAGCCCAAGATGATCGGCGGCATCGCGGCTTCATCA 3505
QY 539 AGTGGCCAGTACGACACATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db AGTGGCCAGTACGACACATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
QY 3506 AGTGGCCAGTACGACACATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3565
Db AGTGGCCAGTACGACACATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3565
QY 599 TGCTGATGCGGCCACCGCCGCTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 658
Db TGCTGATGCGGCCACCGCCGCTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 658
QY 3566 TGCTGTGGGCCACCGCCGCTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 3625
Db TGCTGTGGGCCACCGCCGCTGAACATCATCGCGCCCAACATGCTGACCCAGCTGGGCT 3625
QY 659 GCACCTGAATCTCCCATCAGCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCA 718
Db GCACCTGAATCTCCCATCAGCCCATCGAGACCGTGGCCGCTGAAGCTGAAGCCCGGCA 3685
QY 719 TGGACGCCCCCAAGGTCAGCACTGGCCCTGACCCGAGGAGAGATCAAGGCCCTGACCG 778
Db TGGACGCCCCCAAGGTCAGCACTGGCCCTGACCCGAGGAGAGATCAAGGCCCTGACCG 778
QY 3686 TGGACGCCCCCAAGGTCAGCACTGGCCCTGACCCGAGGAGAGATCAAGGCCCTGGTGG 3745
Db TGGACGCCCCCAAGGTCAGCACTGGCCCTGACCCGAGGAGAGATCAAGGCCCTGGTGG 3745
QY 779 CCATCTCGGAGGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCGCCGAGAACCCCT 838
Db CCATCTCGGAGGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCGCCGAGAACCCCT 838
QY 3746 AGATCTGCACCGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCGCCGAGAACCCCT 3805
Db AGATCTGCACCGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCGCCGAGAACCCCT 3805

QY 839 ACAACACCCCGCTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
Db ACAACACCCCGCTGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 3865
QY 899 ACTTCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGAGAGTGCAGTGGGCAATCCCC 958
Db ACTTCGCGAGCTGAACAAAGCGCACCCAGGACTTCTGGAGAGTGCAGTGGGCAATCCCC 3925
QY 959 ACCCGCGCGGCTGAAGAAAGAGAGACGCTGACCGTCTGGACGTGGGCGAGCGCTACT 1018
Db ACCCGCGCGGCTGAAGAAAGAGAGACGCTGACCGTCTGGACGTGGGCGAGCGCTACT 3985
QY 1019 TCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACACGCTTCACCATCCCCAGCATCA 1078
Db TCAGCTGCCCTGGACGAGGACTTCCGCAAGTACACACGCTTCACCATCCCCAGCATCA 4045
QY 1079 ACAACGAGACCCCGGATCCGCTACCAAGTACAAGTCTGCCCGAGGCTGGAAAGGCA 1138
Db ACAACGAGACCCCGGATCCGCTACCAAGTACAAGTCTGCCCGAGGCTGGAAAGGCA 4105
QY 1139 GCGCCAGCATCTTCAGAGCAGCATGACCAAGATCTCGAGCGCTTCCGCGCGCGCAACC 1198
Db GCGCCAGCATCTTCAGAGTGCAGCATGACCAAGATCTCGAGCGCTTCCGCGAGCAACC 4165
QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTACGTGGGCGAGCGCTGGAGATCG 1252
Db CCGAGATCGTATCTACCAAGTACATGGAGGACCTGTACGTGGGCGAGCGCTGGAGATCG 4225
QY 1253 GCCAGCACCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 1312
Db GCCAGCACCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 4285
QY 4226 GCCAGCACCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGCTTCACCA 4345
QY 1313 CCGCAAGTGGACCGTGCAGCGCCATCGAGCTGCGGAGAGGAGCGCTGACGCTGACAG 1432
Db CCGCAAGTGGACCGTGCAGCGCCATCGAGCTGCGGAGAGGAGCGCTGACGCTGACAG 4405
QY 1433 ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGCAGCAGATCTACCCCGCATCAAGG 1492
Db ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGCAGCAGATCTACCCCGCATCAAGG 4465
QY 1493 TGGCGCAGCTGTCAAGCTGCTGCGCGCGCCAGGCGCTGACCCAGCATCTGCCCTTGA 1552
Db TGGCGCAGCTGTCAAGCTGCTGCGCGCGCCAGGCGCTGACCCAGCATCTGCCCTTGA 4525
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCGCGAGAACCGCGAGATCTGCGCGAGCGCGTGCACG 1612
Db CCGAGGAGCGGAGCTGGAGCTGGCGCGAGAACCGCGAGATCTGCGCGAGCGCGTGCACG 4585
QY 1613 CGGTGTACTAGGACCCAGCAGGAGCTGTGGCGCGAGATCCAGAGCAGGCGCCAGCACC 1672
Db CGGTGTACTAGGACCCAGCAGGAGCTGTGGCGCGAGATCCAGAGCAGGCGCCAGCACC 4645
QY 1673 AGTGGACCTACCATCTACCGAGCGCTTCAAGAACCTCAAGACCGCGCAAGTACGCCA 1732
Db AGTGGACCTACCATCTACCGAGCGCTTCAAGAACCTCAAGACCGCGCAAGTACGCCA 4705
QY 1733 AGATGCGCAGCGCCGACCAACAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 1792
Db AGATGCGCAGCGCCGACCAACAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 4765
QY 1793 CCATGGAGAGCATGCTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAAG 1852
Db CCATGGAGAGCATGCTGATCTGGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAAAG 4825
QY 1853 AGACCTGGGAGACCTGGTGGACCGACTACTTGGCAGGCGCCAGCTGGATCCCGAGTGGAGT 1912
Db AGACCTGGGAGGCGCTGGTGGACCGAGTACTTGGCAGGCGCCAGCTGGATCCCGAGTGGAGT 4885
QY 1913 TCCTGAACACCCCGCTGCTGAAGCTGTGGTACCAAGCTGGGAGAGGAGGCCCATCATCG 1972
Db TCCTGAACACCCCGCTGCTGAAGCTGTGGTACCAAGCTGGGAGAGGAGGCCCATCATCG 1972

Db 4886 TCCTGAACACCCCCCTTGTGAAGCTGTGTACCAGCTGGAGAGGAGCCATCATCG 4945
 Qy 1973 GGGCCAGACCTTCTACGTGGAGGGCGCGCCAAACCGCCAGACCAAGATCGGCAAGCGCG 2032
 Db 4946 GGGCCAGACCTTCTACGTGGAGGGCGCGCCAAACCGCCAGACCAAGATCGGCAAGCGCG 5005
 Qy 2033 GCTACCTGACCGAGCGGGCGCGCAGCAAGATCGTGAGCCCTGACCGAGACCAACCAACGAGA 2092
 Db 5006 GCTACCTGACCGAGCGGGCGCGCAGCAAGATCGTGAGCCCTGACCGAGACCAACCAACGAGA 5065
 Qy 2093 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACAGCGGCGAGGTTGAACATCG 2152
 Db 5066 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGACACAGCGGCGGAGGTTGAACATCG 5125
 Qy 2153 TGACCGACAGCAGTACGCGCTGGGCGATCATCCAGGCGCCAGCGCCGACAGAGCGGAGCG 2212
 Db 5126 TGACCGACAGCAGTACGCGCTGGGCGATCATCCAGGCGCCAGCGGCGGAGAGCGGAGCG 5185
 Qy 2213 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGGTGTACCTGAGCTGGG 2272
 Db 5186 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGGTGTACCTGAGCTGGG 5245
 Qy 2273 TGCCCGCCACAGGCGATCGGCGGCAACAGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2332
 Db 5246 TGCCCGCCACAGGCGATCGGCGGCAACAGAGCAGATCGACAAAGCTGGTGAGCGCGGCA 5305
 Qy 2333 TCCGCAAGGTGCTTCCCTGGAGCGCATCGATGGCGCGCATCGTGATCTACCAGTA 2387
 Db 5306 TCCGCAAGGTGCTTCCCTGGAGCGCATCGACAAAGGCCCAAGGAGGAGGAGCGAGAA 5360
 RESULT 12
 ID ABK91613 standard; DNA; 12411 BP.
 AC ABK91613;
 DT 14-AUG-2002 (first entry)
 XX Modified HIV protein-encoding plasmid DNA #165.
 DE
 XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
 XX Synthetic.
 OS
 PN WO200232943-A2.
 XX
 XX 25-APR-2002.
 XX
 XX 14-AUG-2001; 2001WO-US25721.
 XX
 XX 14-AUG-2000; 2000US-225097P.
 PR 14-NOV-2000; 2000US-252115P.
 PR 28-MAR-2001; 2001US-279257P.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (CHAD/) CHADRABARTI B K.
 XX
 XX Nabel GJ, Huang Y;
 PI
 XX
 XX WPI; 2002-452382/48.
 DR
 XX
 XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
 PT its encoded protein, useful as vaccines for genetic or protein
 PT immunisation for acquired immunodeficiency syndrome or human
 PT immunodeficiency virus infection -
 XX
 XX Disclosure; Page 756-760: 794pp; English.
 PS
 XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
 CC genetic or protein immunisation to a host, respectively. In particular
 CC these are useful for ameliorating the symptoms of acquired
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
 CC compositions are useful for treating or preventing HIV infections or
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
 CC plasmid DNA sequences of the invention.
 XX

SQ Sequence 12411 BP; 2938 A; 3733 C; 3476 G; 2264 T; 0 other;

Query Match 81.2%; Score 2000.2; DB 24; Length 12411;
 Best Local Similarity 91.2%; Pred. No. 1e-242;

Matches 2185; Conservative 0; Mismatches 188; Indels 22; Gaps 5;

Qy 14 TGCCCGAGGCGCATGAGCCAGG---CCACAGCGCCCAACATCTCTGTATGACGCGGCAACT 70
 Db 2967 TGCCCGAGGCGCATGAGCCAGGTCACCAACAGCGCCACCATCATGATGACGCGGCAACT 3026
 Qy 71 TCRAAGGGCCCCAAGCGCATCAAGTGTCTCAACTTGCAGCAAGGAGGGCCACATCGCCC 130
 Db 3027 TCCGCAACACAGCCCAAGATCGTGAAGTGTCTCAACTTGCAGCAAGGAGGGCCACACGCC 3086
 Qy 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTCTGGAAGTGCAGCAAGGAGGGCCACACAGA 190
 Db 3087 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTCTGGAAGTGCAGCAAGGAGGGCCACAGA 3146
 Qy 191 TGAAGACTGACCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 250
 Db 3147 TGAAGACTGACCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3205
 Qy 251 GCAAGCGCGCGAGTTTCCCGAGCGAGCAGAAACCGCGCAACAGCCCGCCACAGCGCGCGAGC 310
 Db 3206 GGAAGCGCGAGGAAATTTCTTACAGCAGCAGACAGACAGCCCGCCACAGAGAGAGC 3265
 Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCGAGCGCGCGGCGGCGGCGGCGGCGGCGGCA 364
 Db 3266 TTCAGGTTTGGGAAGAGAGACAACACTCCCTCTCAGAAGCAGGAGCGGATAGACAGGAA 3325
 Qy 365 -----CCCTGAACCTTCCCGAGATCACCTGTGGAGCGCGCCCTTGTGAGCATCAAG 418
 Db 3326 CTGTATCTTCTAGCTTCCCTCAGATCACTCTTGGCAGCAGCCCTCTGTCACTAATAAGA 3395
 Qy 419 TGGGCGCCAGATCAAGGAGGCGCTCTGTGACACCGCGCGCGCGGCGGCGGCGGCGGCGG 478
 Db 3386 TAGGGGCGCAGCTGAGGAGGCGCTTCTAGACACCGCGCGCGGCGGCGGCGGCGGCGG 3445
 Qy 479 AGATGAGCCTTCCCGCGCAAGTGAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 538
 Db 3446 AGATGAACCTGCCCGCGCGCTGGAAGCCCAAGATGATCGGCGGCGATCGGCGGCTTCATCA 3505
 Qy 539 AGGTGCGCCAGTAGCAGCAGATCCTGATCGAGATCTCGGCGCAAGAGCCCATCGGACCG 598
 Db 3506 AGGTGCGCCAGTAGCAGCAGATCCTGATCGAGATCTCGGCGCAAGAGCCCATCGGACCG 3555
 Qy 599 TGCTGATCGGCGCCCGCT 658
 Db 3566 TGCTGATCGGCGCCCGCT 3625
 Qy 659 GCACCTTGAACTTCCCGCATCAGCCCCCATCGAGACCGGTGCGCGGTGAAGCTGAAGCGCGCA 718
 Db 3626 GCACCTTGAACTTCCCGCATCAGCCCCCATCGAGACCGGTGCGCGGTGAAGCTGAAGCGCGCA 3685
 Qy 719 TGACGCGCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
 Db 3686 TGACGCGCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGCTGG 3745
 Qy 779 CCATCTCGAGGAGATGGAGAAGGAGGCGCAAGATCAACAAGATCGGCGCGCGGAGAACCCCT 838
 Db 3746 AGATCTGACCGAGATGGAGAAGGAGGCGCAAGATCAGCAAGATCGGCGCGCGGAGAACCCCT 3805

QY	839	ACAACACCCCGCTGTTCCGCATCAAGAGAGAGGACAGACCAAGTGGCGCAAGCTGGTGG	898
Db	3806	ACAACACCCCGCTGTTTCGCATTCAAGAAGAAGGACAGACCAAGTGGCGCAAGCTGGTGG	3865
QY	899	ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGATCCCCC	958
Db	3866	ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGTGCAGCTGGGATCCCCC	3925
QY	959	ACCCCGCCGCGCTGAAGAAGAAGAGCGTGAACCGTGTGGAGCTGGGAGCGCCTACT	1018
Db	3926	ACCCCGCCGCGCTGAAGCAGAAGAAGAGCGTGAACCGTGTGGAGCTGGGAGCGCCTACT	3985
QY	1019	TCAGCGTGGCCCTGGACGAGGACTTCGCAAGTACACCGCTTTCACATCCCGAGATCA	1078
Db	3986	TCAGCGTGGCCCTGGACGAGGACTTCGCAAGTACACCGCTTTCACATCCCGAGATCA	4045
QY	1079	ACAACGAGACCCCGGCATCCGCTACCAAGTACAACGCTGCTGCCCGAGGGTGGAAAGGCA	1138
Db	4046	ACAACGAGACCCCGGCATCCGCTACCAAGTACAACGCTGCTGCCCGAGGGTGGAAAGGCA	4105
QY	1139	GCCCCAGCATTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGGCCGCAACC	1198
Db	4106	GCCCCAGCATTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGAACGAGAACC	4165
QY	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTGTAGCTGGGAGGACCTGGAGATCG	1252
Db	4166	CCGACATCGTGATCTACCAAGTACATGGAGACCTGTAGCTGGGAGGAGCCTGGAGATCG	4225
QY	1253	GCAGACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTCTGCGTGGGGCTTCACCA	1312
Db	4226	GCAGACCGCACCAAGATCGAGGAGCTGCGGACAGCCTGCTGGCTGGGGCTTCACCA	4285
QY	1313	CCCCGCACAAGACCAACGAAGAGCGCCCTTCTGTGATGGGTACGAGTGCACC	1372
Db	4286	CCCCGCACAAGACCAACGAAGAGCGCCCTTCTGTGATGGGTACGAGTGCACC	4345
QY	1373	CCGACAGTGGACCTGACGCCATCTGAGCTGCGGAGAAGGAGAGCTGACCGTGAACG	1432
Db	4346	CCGACAGTGGACCTGACGCCATCTGCTGCTGCCGAGAAGGAGAGCTGACCGTGAACG	4405
QY	1433	ACATCAGAAGCTGGTGGCAAGCTGAATGGCGCAGCAGATCTACCCGGGATCAAGG	1492
Db	4406	ACATCAGAAGCTGGTGGCAAGCTGAATGGCGCAGCAGATCTACCCGGGATCAAGG	4465
QY	1493	TGCGCCAGCTGTGCAAGCTGCTGCGCGCGGCCAAGGCCCTGACCGACATCGTGCCTCGA	1552
Db	4466	TGCGCCAGCTGTGCAAGCTGCTGCGCGGCCAAGGCCCTGACCGAGTGGTGCCTCGA	4525
QY	1553	CCGAGGAGCGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGGCGAGCCGCTGCACG	1612
Db	4526	CCGAGGAGCGCGAGCTGGAGCTGGCCGAGAACCGCGAGATCTTGAAGGAGCCGCTGCACG	4585
QY	1613	GCCTGTACTACGACCCACGACGAGGACCTGGTGGCGAGATCCGAGACGAGGGCCACGACC	1672
Db	4586	GCCTGTACTACGACCCACGACGAGGACCTGATGGCGAGATCCGAGACGAGGGCCACGACC	4645
QY	1673	AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA	1732
Db	4646	AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA	4705
QY	1733	AGATGCGACCCGCCACACCAACGACGTAAGACAGCTGACCGAGGCGCTGCAGAAAGTGC	1792
Db	4706	GCATGAAGGGCGCCACACCAACGACGTTGAAGAGCTGAACCGAGGCGCTGCAGAAAGTGC	4765
QY	1793	CCATGGAGAGCATCTGTATCTGGGGCAAGACCCCGAAGTTCCGCTTGCATCCAGAAAGG	1852
Db	4766	CCACCGAGAGCATCTGTATCTGGGGCAAGACCCCGAAGTTTCAAGCTGCCATCCAGAAAGG	4825
QY	1853	AGACCTGGGAGACCTGGTGGACCGACTACTGCGAGGCCAAGCTGCATCCCGAGTGGGAGT	1912
Db	4826	AGACCTGGGAGGCTGGTGGACCGAGTACTGCGAGGCCAAGCTGCATCCCGAGTGGGAGT	4885
QY	1913	TCGTGAACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGAAAGGAGCCATCATCG	1972

[illegible]

RESULT 13

RESOLUTION
ABK91624

ABK91624
ID ABK91624 standard; DNA: 9785 BP.

AC ABK91624:

14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #176.

AA HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

AA	Synthetic.
OS	Synthetic.

AA
PN
WO200232943-A2.

25-APR-2002

14-AUG-2001; 2001WO-US25721.

14-AUG-2000: 2000US-225097P.

FK	14-AUG-2000; 2000US-223037P;
PR	14-NOV-2000; 2000US-252115P;

PR 28-MAR-2001; 2001US-279257P.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRABARTI B K.

2	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	4
---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---

	C	6
	V	7
	C	8
	C	9
	C	10
	L	11
	A	12
	C	13
	.	14
	C	15
	C	16
	F	17
	G	18
	T	19
XX	20	21
XX	22	23

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -

PS Disclosure; Page 791-794; 794pp; English.

CC The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

xx
SQ Sequence 9785 BP; 2360 A; 2843 C; 2731 G; 1851 T; 0 other;

Query Match 81.1%; Score 1996.8; DB 24; Length 9785;
Best Local Similarity 91.3%; Pred. No. 2.8e-242;
Matches 2183; Conservative 0; Mismatches 182; Indels 27; Gaps 5;

QY 14 TGGCCGAGCCATAGCCAGCCACACGAGCCCAACATCTCTGATGTCAGCGGAGCAACTTCA 73
Db TGGCCGAGCCATAGCCAGCCACACGAGCCCAACATCTCTGATGTCAGCGGAGCAACTTCA 3026
QY 74 AGGCCCCAAGCGCATCATCAAGTGCCTCAACTGCGCAAGGAGGCCACATCGCCGCA 133
Db AGGCCCCAAGCGCATCATCAAGTGCCTCAACTGCGCAAGGAGGCCACATCGCCGCA 3093
QY 134 ACTGCCGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCGCAAGGAGGCCACACAGATGA 193
Db ACTGCCGCCCCCGCAAGAGGGCTGCTGGAAGTCCGCGCAAGGAGGCCACACAGATGA 3143
QY 194 AGGACTCAGCGAGCCGACGACCAACTCTTCCGCGAGGACCTGGCCCTTCCCCAGGGCA 253
Db AGGACTCAGCGAGCCGACGACCAACTCTTCCGCGAGGACCTGGCCCTTCCCCAGGGCA 3197
QY 254 AGGCCCGCGAGTTCGCCAGCGAGCAAGCGCCACAGCCACAGCCACCGAGCGCTGC 313
Db AGGCCCGCGAGTTCGCCAGCGAGCAAGCGCCACAGCCACAGCCACCGAGCGCTGC 3257
QY 314 AGGTGCGCGG-----CGAACACCCCGCGAGCGAGCGCGCGCGAGCGCGAGGGCA--- 364
Db AGGTGCGCGGAGAGCAACAACCTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGAACTG 3317
QY 365 ---CCCTGAACCTCCCGCAGATCACTCTGGAGAGCCCGCCCTGGTGAGGATCAAGTGG 421
Db TATCTTTTATGCTTCCCTCAGATCACTCTTGGCAGCAGCCCTCTGTCACAAATGAATAG 3377
QY 422 GCGGCCAGATCAAGAGGCGCTCTGTGACACCGCGCGCGAGCACACCGCTGCGAGGAGA 481
Db GCGGCCAGATCAAGAGGCGCTCTGTGACACCGCGCGCGAGCACACCGCTGCGAGGAGA 3437
QY 482 TGAGCCTGCCCGGGAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAGG 541
Db TGAACCTGCCCGGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCAAGG 3497
QY 542 TGCGCCAGTACGACCATCTGATCAGATCTCGCGCAAGAGGCCATCGGCACCTGC 601
Db TGGCCAGTACGACCATCTGATCAGATCTCGCGCAAGAGGCCATCGGCACCTGC 3557
QY 602 TGATCGCGCCACCGCGTGAATCATATCGCGCGCAACATGCTGACCCAGCTGGGCTGCA 661
Db TGGTGGCCACCGCGTGAATCATATCGCGCGCAACATGCTGACCCAGCTGGGCTGCA 3617
QY 662 CCTGAACTTCCCATCAGCCCATCGAGACCGCTGCGCGCTGGAAGCTGAAGCCCGGCAATGG 721
Db CCTGAACTTCCCATCAGCCCATCGAGACCGCTGCGCGCTGGAAGCTGAAGCCCGGCAATGG 3677
QY 722 ACGGCCCAAGGTGAAGCAGTGCCTTGACCGAGGAGAGATCAAGGCCCTGACCCCA 781
Db ACGGCCCAAGGTGAAGCAGTGCCTTGACCGAGGAGAGATCAAGGCCCTGAGGAGA 3737
QY 782 TCTGCGAGGAGTGGAGAGGGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTTACA 841
Db TCTGCGAGGAGTGGAGAGGGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTTACA 3797

QY 842 ACACCCCGCTTTCGCCCATCAAGAAAGAGACAGCACCAAGTGGCGCAAGCTGTGGACT 901
Db ACACCCCGCTTTCGCCCATCAAGAAAGAGAGACAGCACCAAGTGGCGCAAGCTGTGGACT 3857
QY 902 TCCGCGAGCTGAACAAGGACCCAGGACTTCTGGGAGGTGACGTGGGCAATCCCCAC 961
Db TCCGCGAGCTGAACAAGGACCCAGGACTTCTGGGAGGTGACGTGGGCAATCCCCAC 3917
QY 962 CCGCGCGCTGAACAAGAAAGAGCGGTGACCGTCTGGAGCTGGGCGACGCTACTTCA 1021
Db CCGCGCGCTGAACAAGAAAGAGCGGTGACCGTCTGGAGCTGGGCGACGCTACTTCA 3977
QY 1022 GCCTGCCCTTGAGAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAACA 1081
Db GCCTGCCCTTGAGAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCAACA 4037
QY 1082 ACGAGACCCCGGATCCGCTACCACTACACGTGCTGCCCCAGGCTGGAGGGCAGCC 1141
Db ACGAGACCCCGGATCCGCTACCACTACACGTGCTGCCCCAGGCTGGAGGGCAGCC 4097
QY 1142 CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCAACCCG 1201
Db CCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCAACCCG 4157
QY 1202 AGATCGTATCTACCA-----GCCCGCTGTACGTGGGAGCGACCTTGGAGATCGGCC 1255
Db AGATCGTATCTACCA-----GCCCGCTGTACGTGGGAGCGACCTTGGAGATCGGCC 4217
QY 1256 AGCACCGCGCAAGATCGAGGAGCTGGCAAGCACTGCTGCGCTGGGCTTACACCC 1315
Db AGCACCGCGCAAGATCGAGGAGCTGGCAAGCACTGCTGCGCTGGGCTTACACCC 4277
QY 1316 CCGCAAGAAGACCCAGAAAGGCGCCCTTCTCTGTGGTGGGTACGAGCTGCACCCG 1375
Db CCGCAAGAAGACCCAGAAAGGCGCCCTTCTCTGTGGTGGGTACGAGCTGCACCCG 4337
QY 1376 ACAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCCGTGAACGACA 1435
Db ACAGTGGACCGTGCAGCCCATCGTGTGCGGAGAGAGAGCTGGACCCGTGAACGACA 4397
QY 1436 TCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGACCGCAGATACCCCGGCAATCAAGTGC 1495
Db TCCAGAGCTGGTGGGCAAGCTGAACCTGGGCGACCGCAGATACCCCGGCAATCAAGTGC 4457
QY 1496 GCAGCTGTCAAGCTGCTGCGCGCGCAAGGCGCTTCCGCGACATCGTGCCTGACCG 1555
Db GCAGCTGTCAAGCTGCTGCGCGCGCAAGGCGCTTCCGCGACATCGTGCCTGACCG 4517
QY 1556 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACGCG 1615
Db AGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACGCG 4577
QY 1616 TGCTACGACCCCGCAGCAAGCACTGCTGGCGGAGATCCAGAAAGCGGCGCACACCACT 1675
Db TGCTACGACCCCGCAGCAAGCACTGCTGGCGGAGATCCAGAAAGCGGCGCACACCACT 4637
QY 1676 GGACCTACAGATCTACAGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGA 1735
Db GGACCTACAGATCTACAGAGCGCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGA 4697
QY 1736 TGGCGACCGCCCAACAGAGCTGAAGAGCTGACCGAGGCGCTGAGAAAGATCGCCA 1795
Db TGAAGGCGCCCAACAGAGCTGAAGAGCTGACCGAGGCGCTGAGAAAGATCGCCA 4757
QY 1796 TGGAGACATCTGATCTGGGCGAAGACCCCAAGTTCCGCTGCCCATCCAGAAAGAGA 1855
Db TGGAGACATCTGATCTGGGCGAAGACCCCAAGTTCCGCTGCCCATCCAGAAAGAGA 4817
QY 1856 CTTGGGAGACTGCTGGAGCGGACTTACTGGCAGGCGACCTGGATCCCGAGTGGGAGTGC 1915
Db CTTGGGAGACTGCTGGAGCGGACTTACTGGCAGGCGACCTGGATCCCGAGTGGGAGTGC 4877
QY 1916 TGAACACCCCGCTTGGTGAAGCTGTGGTACAGCTGGAGAGAGGCCCATCATCGGCG 1975

Db 4878 TGAACACCCCCCTGGTGAAGCTGGTACCAGCTGGAGAGGAGCCCATCATCGCGG 4937
QY 1976 CCGAGACCTTCTAGCTGGAGCGGCCGCCAACCGCGAGACAGATCGGCAAGCCCGCT 2035
Db 4938 CCGAGACCTTCTAGCTGGAGCGGCCGCCAACCGCGAGACAGATCGGCAAGCCCGCT 4997
QY 2036 ACCTGACCGACCGGGCGCGCAGAAATCGTGAAGCTGACCGAGACCAACCAAGAA 2095
Db 4998 ACCTGACCGACCGGGCGCGCAGAAATCGTGAAGCTGACCGAGACCAACCAAGAA 5057
QY 2096 CCGAGCTGACAGGCCATCCAGCTGGCCCTGCAGGACACGCGCAGAGGTGAACATCGTA 2155
Db 5058 CCGAGCTGACAGGCCATCCAGCTGGCCCTGCAGGACACGCGCTGGAGGTGAACATCGTA 5117
QY 2156 CCGACACAGTAGTACCGCTGGGATCATCAGGCCCGCAGCAAGAGCGAGCGAGC 2215
Db 5118 CCGACACAGTAGTACCGCTGGGATCATCAGGCCCGCAGCAAGAGCGAGCGAGC 5177
QY 2216 TGGTGAACACAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 2275
Db 5178 TGGTGAACACAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGGTGC 5237
QY 2276 CCGCCCAAGAGGCATCGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGATCC 2335
Db 5238 CCGCCCAAGAGGCATCGCGGCAACGAGCAGGTGGAGCGCTGGTGAGCGCGCGCATCC 5297
QY 2336 GCAAGTGTCTTCTCGGAGCGCATCGATGGCGGCATCGTGTACAGTA 2387
Db 5298 GCAAGTGTCTTCTCGGAGCGCATCGACAAAGGCCCGAGGAGGACGAGAA 5349

RESULT 14

ABK91619
ID ABK91619 standard; DNA; 9167 BP.
AC ABK91619;
XX
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #171.
XX
DE HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
XX WO200232943-A2.
XX
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001WO-US25721.
XX
XX 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRAHARTI B K.
XX
XX Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX
XX Disclosure; Page 775-778; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 9167 BP; 2225 A; 2668 C; 2517 G; 1757 T; 0 other;

Query Match 80.9%; Score 1993.4; DB 24; Length 9167;
Best Local Similarity 91.1%; Pred. No. 7.6e-242;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;
QY 14 TGGCCGAGGGCCATGAGCCAGG---CCACCAGCGCCACATCCTGATGAGCGCAGCAACT 70
Db 2964 TGGCCGAGGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3023
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCACTGGGGCAAGAGGGCCACATTCGCC 130
Db 3024 TCCGCAACCCAGCGCAAGATCGTGAAGTCTTCAACTTGGGCAAGAGGGCCACACCGCC 3083
QY 131 GCAACTCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACCA 190
Db 3084 GCAACTCGCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACCA 3143
QY 191 TGAAGGAGTGCACGAGCGCCAGCGCCAACTTCTCCGCGAGGAGCTGGCCCTTCCGCCAGG 250
Db 3144 TGAAGGAGTGCACGAGCGCACAGGCTAA-----TAGGGAAGATCTGGCCCTTCCACAAG 3197
QY 251 GCAAGGGCCCGGAGTTTCCCGAGCGAGCAGAGAACCGCGCCCAACAGCCCAACAGCCCGAGC 310
Db 3198 GGAAGGGCAGGGAATTTTCTTCAGAGCAGACAGAGCCACAGAGCCCAACAGAGAGAGC 3257
QY 311 TGCAGGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCAGGGCA 364
Db 3258 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAAAGAGAGGCCGATAGACAAGGAA 3317
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGTGAGCATCAAG 418
Db 3318 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCAACAATAAGA 3377
QY 419 TGGCGGCGCAGATCAAGAGGCCCTGTGTGACACCGGGCGCGAGCAGACACCGTGTGGAGG 478
Db 3378 TAGGGGGCCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGAGCACACCGTGTGGAGG 3437
QY 479 AGATGAGCTGCCCGCAAGTGAAGCCCAAGATGATCGCGGGCATCGCGGCTTCATCA 538
Db 3438 AGATGAACCTGCCCGCGCTTGAAGCCCAAGATGATCGCGGGCATCGCGGCTTCATCA 3497
QY 539 AGGTGCGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGCCATCGCACCG 598
Db 3498 AGGTGGGCCAGTACGACAGATCTCTGATCGAGATCTGCGGCCACAAGGCCATCGCACCG 3557
QY 599 TGCTGATGGGCCACCGCCCTGAAACATCATCGGCCCGCAACATGCTGACCGAGCTGGGCT 658
Db 3558 TGCTGTTGGGCCCGCCCGCTGAAACATCATCGGCCCGCAACATGCTGCTACCGAGATCGGCT 3617
QY 659 GCACCTTGAACCTTCCCATCATCAGCCCATCGAGACCGTCCCGTGAAGCTGAAGCCCGGCA 718
Db 3618 GCACCTTGAACCTTCCCATCATCAGCCCATCGAGACCGTCCCGTGAAGCTGAAGCCCGGCA 3677
QY 719 TGGAGCGCCCCAAGGTGAAGCAGTGGCCCTTCACCGAGGAGAAATCAAGCCCTGACCG 778
Db 3678 TGGAGCGCCCCAAGGTGAAGCAGTGGCCCTTCACCGAGGAGAAATCAAGCCCTGCTGGTGG 3737
QY 779 CCATCTCGGAGGAGATGGAGAGGAGGCAAGATCAACAGATCGGCCCGCCGAGAACCCCT 838
Db 3738 AGATCTGACCGAGATGGAGAGGAGGCAAGATCAAGCAAGATCGGCCCGCCGAGAACCCCT 3797

QY 839 ACAACACCCCGTTCGCGCATCAAGAAAGAGACACCAAGTGCAGCAAGCTGTGG 898
Db 3798 ACAACACCCCGTTCGCGCATCAAGAAAGAGACACCAAGTGCAGCAAGCTGTGG 3857
QY 899 ACTTCCGGAGCTGAACAAGCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958
Db 3858 ACTTCCGGAGCTGAACAAGCACCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC 3917
QY 959 ACCCGCGCGCTGAAGAAGAAGAGCGTGACCGTGTCTGGAGCTGGGCGACGCCCTACT 1018
Db 3918 ACCCGCGCGCTGAACAGAGAGAGCGTGACCGTGTCTGGAGCTGGGCGACGCCCTACT 3977
QY 1019 TCAGCTGCCCTTGGAGAGAGACTTCCGCAAGTACACCGCCTTCAACATCCCAAGCATCA 1078
Db 3978 TCAGCTGCCCTTGGAGAGAGACTTCCGCAAGTACACCGCCTTCAACATCCCAAGCATCA 4037
QY 1079 ACAACGAGACCCCGGATCCGCTACCAATACACGCTGCTGCCCGAGGCTGGAAGGCA 1138
Db 4038 ACAACGAGACCCCGGATCCGCTACCAATACACGCTGCTGCCCGAGGCTGGAAGGCA 4097
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCCTGGAGCCCTTCCGGCCCGCAACC 1198
Db 4098 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCCCTGGAGCCCTTCCGGCAAGCAACC 4157
QY 1199 CGAGATCTGATCTACCA-----GGCCCCCTGTACGTGGGAGCAGCTGGAGATCG 1252
Db 4158 CGAGATCTGATCTACCAATGATGACCAACCTGTACGTGGGAGCAGCTGGAGATCG 4217
QY 1253 GCAGACCGCGCAAGATCGAGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTACCA 1312
Db 4218 GCAGACCGCGCAAGATCGAGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTACCA 4277
QY 1313 CCGCCGCAAGAGACACCAAGAGAGCCCGCTTCTGTGTGGTGGCTAGAGCTGCACC 1372
Db 4278 CCGCCGCAAGAGACACCAAGAGAGCCCGCTTCTGTGTGGTGGCTAGAGCTGCACC 4337
QY 1373 CCGACAAGTGGAGCTGCAGCCCATCGAGCTGCCGAGAAGAGAGCTGGACCGTGAACG 1432
Db 4338 CCGACAAGTGGAGCTGCAGCCCATCGAGCTGCCGAGAAGAGAGCTGGACCGTGAACG 4397
QY 1433 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGCATCAAG 1492
Db 4398 ACATCCAGAAGCTGTGGGCAAGCTGAACCTGGCCAGCCAGATCTACCCCGCATCAAG 4457
QY 1493 TGGCCAGCTGTGCAAGCTCTCGCGGCCCAAGGCCCTTGAACGACATCGTGCCTGTA 1552
Db 4458 TGGCCAGCTGTGCAAGCTCTCGCGGCCCAAGGCCCTTGAACGACATCGTGCCTGTA 4517
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGCGGAGCCGTCGACG 1612
Db 4518 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGGAAGAGCCGTCGACG 4577
QY 1613 GCGTGTACTACGACCCAGCAAGGACTGTGTGGCCGAGATCCAGAGCAGGCGCACGACC 1672
Db 4578 GCGTGTACTACGACCCAGCAAGGACTGTGTGGCCGAGATCCAGAGCAGGCGCACGACC 4637
QY 1673 AGTGGACCTACAGATCTACAGAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1732
Db 4638 AGTGGACCTACAGATCTACAGAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCC 4697
QY 1733 AGATGGCACCCGCGCACCAACAGCTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 1792
Db 4698 GCATGAAGGCGCGCACCAACAGCTGAAGCAGCTGACCGAGCGCGTGCAGAGATCG 4757
QY 1793 CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1852
Db 4758 CCACGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAGG 4817
QY 1853 AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCCACTTGGATCCCGAGTGGGAGT 1912
Db 4818 AGACCTGGGAGACCTGGTGACCGAGTACTTGGCAGGCCACTTGGATCCCGAGTGGGAGT 4877
QY 1913 TCGTGAACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGCCCATCATCG 1972

Db 4878 TCGTGAACACCCCGCTGGTGAAGCTGTGGTACCAGCTGGAAGAGAGCCCATCATCG 4937
QY 1973 GCGCCGAGAGACTTCTAGCTGGACGGCGCCCAACCGGAGAGCAAGATCGGCAAGGCCG 2032
Db 4938 GCGCCGAGAGACTTCTAGCTGGACGGCGCCCAACCGGAGAGCAAGCTGGCAAGGCCG 4997
QY 2033 GCTACGTGACCGACCGGGCGCGGAGAGATCGTGTGAGCTTACCGGAGACCAACCAACGAGA 2092
Db 4998 GCTACGTGACCGACCGGGCGCGGAGAGTGTGTGCCCTGACCGACACCAACCAACGAGA 5057
QY 2093 AGACCGAGCTGACAGGCCATCCAGCTGGCCCTGTCCAGGACAGCGGAGAGTGAACATCG 2152
Db 5058 AGACCGAGCTGACAGGCCATCCAGCTGGCCCTGTCCAGGACAGCGGCTGTGAGGTGAACATCG 5117
QY 2153 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCGCCAGCGGACAGAGCGAGAGCG 2212
Db 5118 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCGCCAGCGGACAGAGCGAGAGCG 5177
QY 2213 AGCTGTGTAACAGAGATCATCGAGCAGCTGTATCAAGAAGAGAGTGTACTGTAGCTGGG 2272
Db 5178 AGCTGTGTAACAGAGATCATCGAGCAGCTGTATCAAGAAGAGAGTGTACTGTAGCTGGG 5237
QY 2273 TGCCCGCCCAACAAGGGCATCGCGGCCAACGACGATCGACAGCTGTGTGAGCAAGGGCA 2332
Db 5238 TGCCCGCCCAACAAGGGCATCGCGGCCAACGACGATCGACAGCTGTGTGAGCGCCGCA 5297
QY 2333 TCAGCAAGTGTCTTCTTCTGACGCGCATCGATGCGCGCATCTGTATCCAGTA 2387
Db 5298 TCAGCAAGTGTCTTCTTCTGACGCGCATCGACAAAGGCCAGGAGGAGACGAGAA 5352

RESULT 15

ABK91614

ID ABK91614 standard; DNA; 9170 BP.

XX ABK91614;

XX 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #166.

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX Synthetic.

XX WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.

PR 28-MAR-2001; 2001US-279257P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -

XX Disclosure; Page 760-763; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

xx
SQ Sequence 9170 BP; 2225 A; 2669 C; 2519 G; 1757 T; 0 other;

Query Match 80.9%; Score 1993.4; DB 24; Length 9170;
Best Local Similarity 91.1%; Pred. No. 7.6e-242;
Matches 2182; Conservative 0; Mismatches 186; Indels 27; Gaps 5;

QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTCTGATGAGCGGAGCAACT 70
DB 2967 TGGCCGAGGCGCATGAGCCAGGCTGACACAGCGCCACCATCATGATGAGCGGCGCAACT 3026
QY 71 TCAAGGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGCAAGGAGGGCCACATCGGCC 130
DB 3027 TCCGCAACCAGCGCAAGATCTGTAAGTCTTCAACTGCGCAAGGAGGGCCACACCGGCC 3086
QY 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
DB 3087 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGCACGAGCGCGCAGGCCACTTCTTCGCGAGGAGCTGGCTTCCCCCAGG 250
DB 3147 TGAAGGACTGCACGAGCGCGCAGGCCACTTCTTCGCGAGGAGCTGGCTTCCCCCAGG 3200
QY 251 GCAAGGGCCGCGAGTCTCCCGAGCGAGCAGAACCGCGCCACAGCCGCCAGCGCGAGC 310
DB 3201 GGAAGGCCAGGATTTCTTCAGAGCAGACCGAGCCACAGCCGCCACAGAGAGAGCG 3260
QY 311 TGCAGTGGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGAGCGGCA 364
DB 3261 TTCAGGTTTGGGGAGAGAGACAACAACTCCCTCTCAGAGCAGGAGCGGATAGACAAGGAA 3320
QY 365 -----CCCTGAACCTCCCGACAGATCACTCTGGCAGCGCCCGCTGGTAGCATCAAGG 418
DB 3321 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCTGTCACAATAAGA 3380
QY 419 TGGCGGCCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGAGCGCGCGCGAGCGCTGGAGG 478
DB 3381 TAGGGGCCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGAGCAGACCGTGTGGAGG 3440
QY 479 AGATGAGCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGCGATCGGGGGCTTCATCA 538
DB 3441 AGATGAACCTGCCCGCGCGCTGGAAGGCCCAAGATGATCGGCGGCGATCGGGGGCTTCATCA 3500
QY 539 AGTGGCGCCAGTAGCAGCATCTGATCGAGATCTGGGCGCAGCAAGGCCATCGGCACCG 598
DB 3501 AGTGGCGCCAGTAGCAGCATCTGATCGAGATCTGGGCGCAGCAAGGCCATCGGCACCG 3560
QY 599 TGTGTATCGCGCCCGCCCGTGAACATCATCGGCGCCCAACATGCTGACCGCAGCTGGGCT 658
DB 3561 TGTGTGTGGGCCCGCCCGCTGATCGAGATCTGGGCGCAGCAAGGCCATCGGCACCGCT 3620
QY 659 GCACCTGTAACCTTCCCATCAGCGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCGCGGCA 718
DB 3621 GCACCTGTAACCTTCCCATCAGCGCCCATCGAGACCGTGGCCCGTGAAGCTGAAGCGCGGCA 3680
QY 719 TGGAGCGCCCGCAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 3681 TGGAGCGCCCGCAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGTTGG 3740
QY 779 CCATCTCGAGGAGATGGAGAAGGAGGCAAGATCAACAAGATCGGCGCCCGCAGAACCCCT 838
DB 3741 AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAACAAGATCGGCGCCCGCAGAACCCCT 3800

QY 839 ACAACACCCCGTGTTCGCCATCAAGAAAGAGCAGACCAACGCTGGCGAAGCTGGTGG 898
DB 3801 ACAACACCCCGTGTTCGCCATCAAGAAAGAGCAGACCAACGCTGGCGAAGCTGGTGG 3860
QY 899 ACTTCCGCGAGCTGAACAAGCGCAGCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCC 958
DB 3861 ACTTCCGCGAGCTGAACAAGCGCAGCCAGGACTTCTGGAGGTGAGCTGGGCGATCCCC 3920
QY 959 ACCCGCGCGGCTGAAGAAGAGAGAGCGTCAACGCTGCTGGACCTGGCGACGCGCTACT 1018
DB 3921 ACCCGCGCGGCTGAAGCAGAGAGAGCGTCAACGCTGCTGGACCTGGCGACGCGCTACT 3980
QY 1019 TCAGCGTGCCTTGSACGAGGACTTCCGCAAGTACACCGCCCTTCACTATCCCCAGCATCA 1078
DB 3981 TCAGCGTGCCTTGSACGAGGACTTCCGCAAGTACACCGCCCTTCACTATCCCCAGCATCA 4040
QY 1079 ACACGAGACCCCGGCTACCGTACCAAGTACAGAGTCTGCCCCAGGCTGGGAAGGGA 1138
DB 4041 ACACGAGACCCCGGCTACCGTACCAAGTACAGAGTCTGCCCCAGGCTGGGAAGGGA 4100
QY 1139 GCCCGAGCATCTCCAGAGCAGCATGACCAAGATCCTTGAGGCCCTTCCGCGCCCGCAACC 1198
DB 4101 GCCCGGCGCTTCCAGTGCAGCATGACCAAGTCTTGAGGCCCTTCCGCAAGAGAAC 4160
QY 1199 CCGAGATCGTGTATACCA-----GGCCCCCTGTACGTGGGCGAGCAGCTGGAGATCG 1252
DB 4161 CCGACATCGTGTATACCAAGTACATGGACACCTGTACGTGGGCGAGCAGCTGGAGATCG 4220
QY 1253 GCCAGCACCGCGCAGAGATCGAGGAGCTGGCGAAGCAGCTTGTGGCTTGGGCTTCACCA 1312
DB 4221 GCCAGCACCGCAGAGATCGAGGAGCTGGCGAAGCAGCTTGTGGCTTGGGCTTCACCA 4280
QY 1313 CCCCCGCAAGAAGCAGCAGAGGAGGCCCCCTTCTGTGGATGGCTACGAGCTGCACC 1372
DB 4281 CCCCCGCAAGAAGCAGCAGAGGAGGCCCCCTTCTGTGGATGGCTACGAGCTGCACC 4340
QY 1373 CCGACAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1432
DB 4341 CCGACAAGTGGACCGTGCAGCCCATCGTGTGCCGAGAGAGCAGCTGGACCGTGAACG 4400
QY 1433 ACATCCAGAACTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAG 1492
DB 4401 ACATCCAGAACTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAG 4460
QY 1493 TGGCGCAGCTGTCAAGCTGTGCGCGCGCAGAGCCCTGACCGACATCTGTGCCCTGA 1552
DB 4461 TGGCGCAGCTGTCAAGCTGTGCGCGCGCAGAGCCCTGACCGAGGTGTGTGCCCTGA 4520
QY 1553 CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTGCGGAGCGCCGTGCACG 1612
DB 4521 CCGAGAGGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTGCGGAGCGCCGTGCACG 4580
QY 1613 GCGTGTACTAGCACCCCGAGAGACCTGTGGTGGCGGAGATCCAGAGCAGGCGCCACGACC 1672
DB 4581 GCGTGTACTAGCACCCCGAGAGACCTGTGGTGGCGGAGATCCAGAGCAGGCGCCACGACC 4640
QY 1673 AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAAGCTGAAGACCGCGAAGTACGCA 1732
DB 4641 AGTGGACCTACAGATCTACCGAGGCGCTTCAAGAAGCTGAAGACCGCGAAGTACGCA 4700
QY 1733 AGATGCGACCGCCCGCACACCGACCTGAGAGCAGCTGACCGAGGCGCGTGCAGAGATCG 1792
DB 4701 GCATGAAGGCGCGCCACACCAACGAGCTGAAGAGCTGACCGAGGCGCGTGCAGAGATCG 4760
QY 1793 CCATGGAGAGCATCGTGTATCTGGGCGAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 1852
DB 4761 CCACCGAGAGCATCGTGTATCTGGGCGAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 4820
QY 1853 AGACTGGGAGCTGTGGAGCGGACTTACGGAGGCGCCACCTGGATCCCCCAGTGGGAGT 1912
DB 4821 AGACTGGGAGCTGTGGAGCGGACTTACGGAGGCGCCACCTGGATCCCCCAGTGGGAGT 4880
QY 1913 TCGTGAACACCCCGCTTGTGAAGCTGTGGTACCGAGCTGGAGAGGAGCGCCATCATCG 1972

```
|||||
Db 4881 TCGTGAACACCCCCCTGGTGAAGCTGTGTACAGCTGGAGAAAGGAGGCCCATCATCG 4940
QY 1973 GCGCCGAGACCTTCTACGTGGAGGGCGCCCAACCCGCGAGACCAAGATCGGCAAGGCCG 2032
Db 4941 GCGCCGAGACCTTCTACGTGGAGGGCGCCCAACCCGCGAGACCAAGCTGGGCAAGGCCG 5000
QY 2033 GCTACGTGACCGACCGGGCGCGGAGAAAGATCGTGAAGCTGACCGCTGACCGACCAACCAACGAGA 2092
Db 5001 GCTACGTGACCGACCGGGCGCGGAGAAAGATCGTGAAGCTGACCGCTGACCGACCAACCAACGAGA 5060
QY 2093 AGACCGAGCTGCAGGGCCATCCAGCTGGCCCTGCAGGACACAGCGCAGCGAGGTGAACATCG 2152
Db 5061 AGACCGAGCTGCAGGGCCATCCAGCTGGCCCTGCAGGACACAGCGCCTGGAGGTGAACATCG 5120
QY 2153 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCCAAGCCGACAGAGCGAGAGCG 2212
Db 5121 TGACCGACAGCCAGTAGCCCTGGGCATCATCCAGGCCCAAGCCGACAGAGCGAGAGCG 5180
QY 2213 AGCTGGTGAACACAGATCATCGACGCTGATCAAGAGGAGAAAGGTGTACCTGAGCTGGG 2272
Db 5181 AGCTGGTGAACACAGATCATCGACGCTGATCAAGAGGAGAAAGGTGTACCTGGCCTGGG 5240
QY 2273 TGCCCGCCCAACAAGGGCATCGGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA 2332
Db 5241 TGCCCGCCCAACAAGGGCATCGGCGGCAACGAGCAGGTGGAGCGCCTGGTGAGCGCGCGCA 5300
QY 2333 TCCGCAAGGTGCTGTTCTTCTGGAGGGCATCGATGGCGGCATCGTGTGATCTACAGTA 2387
Db 5301 TCCGCAAGGTGCTGTTCTTCTGGAGGGCATCGACAAAGGCCCAAGGAGGAGCACAGAGAA 5355
```

Search completed: February 10, 2003, 12:44:09
Job time : 583 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 11:44:48 ; Search time 3386 Seconds
(without alignments)
11780.710 Million cell updates/sec

Title: US-09-610-313-31

Perfect score: 2463

Sequence: 1 gtcagccaccatggccga.....gggctagcaccggtaattc 2463

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estl:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.8	3.5	2598	11 AY103647	AY103647 Zea mays
2	85.6	3.5	951	13 BM321451	BM321451 rockefell
3	80.2	3.3	1132	13 BM320864	BM320864 rockefell
4	79.6	3.2	1165	13 BM320900	BM320900 rockefell
5	79	3.2	867	13 BM321430	BM321430 rockefell
6	76.6	3.1	1550	13 BM321022	BM321022 rockefell

7	75	3.0	862	13	BM321023	rockefell
8	73.2	3.0	853	13	BM321393	rockefell
9	72.8	3.0	566	13	BM587428	BM587428 170006873
10	72.2	2.9	640	10	BE601575	BE601575 HVSME009
11	71.6	2.9	500	13	BM372120	BM372120 EBR003_SQ
12	71.6	2.9	538	13	BM368580	BM368580 EBR003_SQ
13	71.6	2.9	540	9	AJ471121	AJ471121 AJ471121
14	71.6	2.9	566	14	BQ464692	BQ464692 HF02P20r
15	71.6	2.9	578	13	BM377112	BM377112 EBR005_SQ
16	71.6	2.9	579	12	BF253640	BF253640 HVSMEF000
17	71.6	2.9	582	12	BM372871	BM372871 EBR004_SQ
18	71.6	2.9	588	14	BQ765498	BQ765498 EBR003_SQ
19	71.6	2.9	600	9	AL508778	AL508778 AL508778
20	71.6	2.9	625	14	BQ768295	BQ768295 EBR008_SQ
21	71.6	2.9	656	13	BM370222	BM370222 EBR009_SQ
22	71.6	2.9	660	14	BQ762385	BQ762385 EBR001_SQ
23	71.6	2.9	700	9	AL508513	AL508513 AL508513
24	71.6	2.9	768	12	BG365317	BG365317 HVSME1000
25	71.6	2.9	789	12	BF626988	BF626988 HVSMEB000
26	71.6	2.9	822	12	BF417596	BF417596 HVSMEF001
27	71.4	2.9	730	12	BF260696	BF260696 HVSMEF002
28	71.2	2.9	646	13	BM645449	BM645449 170006873
29	71.2	2.9	724	13	BM584191	BM584191 170006872
30	71.2	2.9	1460	11	AY104825	AY104825 Zea mays
31	71.2	2.9	2129	11	AY109914	AY109914 Zea mays
32	71	2.9	925	17	CNS0091P	AL053013 Drosophila
33	70.8	2.9	2299	11	AY106831	AY106831 Zea mays
34	70.6	2.9	861	10	BE636696	BE636696 rockefell
35	70.2	2.9	3134	11	AY109500	AY109500 Zea mays
36	70	2.8	494	12	BG349122	BG349122 947029E04
37	70	2.8	571	12	BF266979	BF266979 HV_CEA001
38	70	2.8	609	13	BJ292939	BJ292939 BJ292939
39	70	2.8	613	13	BJ226790	BJ226790 BJ226790
40	70	2.8	616	13	BJ280623	BJ280623 BJ280623
41	70	2.8	619	13	BJ291714	BJ291714 BJ291714
42	70	2.8	627	12	BG300932	BG300932 HVSMEB001
43	70	2.8	635	12	BJ290451	BJ290451 BJ290451
44	70	2.8	643	13	BJ289971	BJ289971 BJ289971
45	70	2.8	652	13	BJ290578	BJ290578 BJ290578

ALIGNMENTS

RESULT	1	2598 bp	linear	HTC 25-MAY-2002
AY103647	AY103647	PCO142084	mRNA	
LOCUS	Zea mays			
DEFINITION	Zea mays			
ACCESSION	AY103647			
VERSION	AY103647.1	GI:21206725		
KEYWORDS	HTC			
SOURCE	Zea mays			
ORGANISM	Zea mays			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC			
AUTHORS	Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.			
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes			
JOURNAL	Unpublished (2002)			
REFERENCE	2 (bases 1 to 2598)			
AUTHORS	Coe, E.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA			
FEATURES	Location/Qualifiers			
source	1..2598			
	/organism="Zea mays"			
	/db_xref="MaizeDB:638378"			
	/db_xref="taxon:4577"			
	/clone="PCO142084"			

/Clone_lib="Maize Mapping Project/DuPont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project."

BASE COUNT 490 a 907 c 795 g 406 t
ORIGIN

Query Match 3.5%; Score 86.8; DB 11; Length 2598;

Best Local Similarity 42.1%; Pred. No. 6.9e-05;

Matches 813; Conservative 0; Mismatches 1102; Indels 15; Gaps 5;

```

QY 296  CACACAGCCGAGCTGCAGTCCGCGGCGACAAACCCCGCACAGGCGCGCGCGAGC 355
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 51  CCACCACTCCCACTCTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 110
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 356  GCCAGGCGACCCCTGAATTCCTCCCGACATACCTGTGGCAGCGCCCGCTGGTGAGCATCA 415
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 111  TGGCGGCGACGCGGCATCTGTGGAGAGCACCGCTGACTGGGCGCGCGCGCGCG 170
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 416  AGGTGGGCGCGCGAGATCAAGAGGCGCTGTGACACCGCGCGCGACACCGTGTCTGG 475
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 171  AGCTGGCGGCGGAGCCACTTGACGAGGTGAAGCGCATGTTGGCGCAGCGCGCGCGCG 230
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 476  AGGAGATGAGCTTCGCGCGCAAGTGAAGCCCAAGATGATGCGGCGCGCGCGCTTCA 535
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 231  TGTCAAGATTCGAGGCTTCACCTCGCGTTCGCGCGAGTGGCGCGCGCTTCGCGCGCA 290
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 536  TCAAGGTGCGCCAGTACGACAGCATCTGATCGAGATCTCGCGCAAGAGCGCATCGGCA 595
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 291  AGACGCGTTCGCGGTCGCGTTCGAGCTCGACGAGGCGCGCGCGCGCTCAAGGCCA 350
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 596  CCGTGTGATGCGCGCGCGCGCGCGTGAACATATCGCGCGCGCGCGCGCGCGCGCGCG 655
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 351  GCAGCGAGTGGATCTCGACTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 410
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 656  GCTGCGACCTGAATTCCTCCATCAGCCCGCATCGAGCGGTGCGCGTGAAGTGAAGCGCG 715
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 411  GCTTGGCGGCGACTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 716  GCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 775
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 471  TCAGGATCTCAAGCGCGGAATCTTCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 530
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 776  CCGCGATCTCGGAGGATGGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 835
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 531  AGGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 836  CTTACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 895
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 591  TCCGCTTCGAGATCTCGAGGCGCATCACGAGGTGTCTCAACACCGCGTGTCCGCGCGCG 650
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 896  TGGACTTCGCGCGAGCTGAACAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 955
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 651  TGGCGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 710
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 956  CCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1015
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 711  GCCTCATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 767
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1016  ACTTCAGCTGCCCTGGACGAGGACTTCGCGAAGTACACCGCGCTTCACCATCCCCAGCA 1075
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 768  ACAGCGCGGAGGCGTTCAAGATCGCGGCGCATCGAGGCGCGCTTCTTCAAGCTCAACCGCA 827
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1076  TCAACACGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1135
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 828  AGGAGGCGCTCGCGCATCGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 887
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1136  GCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1195
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

DB 888  TGTACGACGCCAACGTCCTGGCGCTCTGTGCGAGGTCTCT--GTCCGCGCTTCTTGCG 944
QY 1196  ACCCGGAGATCGTGATCTACCGAGCGCCCTGTACGTGGCGACGACCTGAGATCGGCC 1255
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 945  AGTCTATGAACGCGCAAGCCCGGATACAGGACCACTTGACCCACACAGTGAAGCACACC 1004
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1256  AGCACCGCGCGCAAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCAACACCC 1315
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1005  CGGGTCCATTCGAGCGCGCGCCATCATGGAGCACATCTCTGGATGGCAGCTCTTCATGA 1064
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1316  CCACAAAGAACACACAGAGGAGCGCCCTCTCTGTGGATGGGTACGAGCTGACACCCCG 1375
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1065  AGCAGGCCAAGAAAGTTGAACGAGCTGGACCCCGCTGTCTGAAGCCCAAGCAGCAGGTAC 1124
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1376  ACAAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCGTGAACACCA 1435
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1125  CGTTCGCGAGTCCGCGAGTGGCTGGGCCCGCCAGATCGAGGTG--ATCCGCGCGCCCA 1181
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1436  TCCAGAGGCTGGTGGCAAGCTGAACCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1495
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1182  CCAAGTCCATCGAGCGCGGAGTCAACTCCGTGAAGCA--CAACCGGTTCATCGACGTCC 1238
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1496  GCCAGCTGTGCAAGCTGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1555
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1239  ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1298
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1556  AGGAGCGCGAGCTGGAGCTGGCGCGAGAACCGGAGATCTCTGCGGAGCGCGCGCGCGCG 1615
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1299  ACAAGCGCGCGCTCGCGCATCGCAACATCGCAAGCTCATGTTGCGCGAGTTCCTCCGAGC 1358
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1616  TGTACTACGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1675
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1359  TCGTCAAGGAGTTCTACAACAACCGGCTCACTCCACCTGGCGCGCGCGCGCGCGCGCGCG 1418
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1676  GGACCTACAGATCTACCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1735
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1419  GCCTGAGTACGCGCTTCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1478
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1736  TGGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1795
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1479  AGTACCTGGCGCACCGCATCAACACCGTGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1538
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1796  TGGAGGAGCAT--CGTGAATCGGGCGCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1852
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1539  TGAATTCCTGGCGCTCGTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1598
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1853  AGACTGGGAGACCTGGTGACCGACTACTGGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1912
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1599  TCATGTGCTCCACTACATCGTGGCGCTGTGCCAGCGCGTGGAGCTTGGCGCGCGCGCGCG 1658
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1913  TCGTGAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1972
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1659  AGACATCAAGGCGTGGTGAAGAACACCGTGACCCAGGTGGCGCGCGCGCGCGCGCGCGCG 1718
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1973  CGCGCGGAGACCTTCTAGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2032
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1719  TGAACCGCTCGGCGAGCTCTCCAGCGCGCGCGTTCACGAGAGAGGAGCTGATCAGCGCGCA 1778
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2033  GCTACGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2092
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1779  TCAGCGCGGAGGCGCTTTCACCTACCGGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1838
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2093  AGACCGAGCTGCGAGGCGCATCCAGCTGGCCCTGCGAGGACAGCGCGCGCGCGCGCGCGCG 2152
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1839  TGCAGAGCTTGGCGCGCGCTGCTGGTGAACACCGCGCTCAGCAGGCGCGCGCGCGCGCGCG 1898
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2153  TGACCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2212
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1899  AGCCCTCCGTTCTCCAAGATCACAGGTTTCGAGGAGGAGCTCCGCGCGCGTGTGCGCC 1958
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 2213  AGCTGGTGA 2222
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1959  AGGAGGTGA 1968
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```



```
Db 253 GAGCTACCCGGCTTCGGCGTCAAGCTCGGCTGACCAACTAGCGCGCGCTACGGCGACT 312
Qy 379 CAGATCACTTGTGGAGCGCCCTGGTGAGCATCAAGTGGCGCGCGAGATCAAGGAG 438
Db 313 GGCTGTGTGGCGCCCGCTGTGCTGAAGAAGCTCAACCTCGACTCCAAGTACGAGGT 372
Qy 439 GCGCTGTGGACACCGCGCGGAGACACCTGCTGAGGAGATGAGCCTGCCGCGCAAG 498
Db 373 GTCAAGAGGTCAACGGCGAGGACTACAAGTTCGAGAGTTCGACGAGCGCGCCCGCT 432
Qy 499 TGAAGCCCAAGATGATCGCGCGCATCGCGGCTTCAAGTGGCGCGAGTACGACGAG 558
Db 433 TTCAAGGCCCTGTCTGACGCTGGCGCTGTTCGCGACCTCGACTGCGCGCGCTTCGCC 492
Qy 559 ATCTGTATGATGATCTCGCGCAGAAGGCGATCGGCACCGTGTGATCGCGCGCCACCC 618
Db 493 GCGCTCAAGGGCATGTGCGGCGCGGCTCAAGCTGCC---CACAGCGAGACCGGCTTC 549
Qy 619 GTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGACCCCTGAACCTCCCATC 678
Db 550 GTCGGCTTCAACGGCGACAGAGGAGCTCAAGCCCGGCTCTTCGCAAGTACATCTTC 609
Qy 679 AGCCCCATGAGACCGTGGCGGCTGAAGCTGAAGCCCGGCGATGAGCGCGCCCAAGGTGAAG 738
Db 610 GCGGCGCACGCTCGCGCGGTATCATGAAGCTCCTCAAGGAGCAGGAGCGCGCGCTTCGAC 669
Qy 739 CAGTGGCCCCGTACCGGAGGAGATCAAGCCCTGACCCCATCTCGCGAGGAGATGGAG 798
Db 670 CGCGAGTTCGCGGCTACGCCAAGGAGGTGTACCGCGGACATGCTCGAGAAGATCTAC 729
Qy 799 AAGGAGGCAAGATCAACCAAGATCGCGCGGAGAACCCCTACACACCGCCCGGTTCCGCC 858
Db 730 ACCGAGCCACAGAGAGATCCGCGGAGCCGACCTTCGTCGCCAAGCGCGCTCGAAG 789
Qy 859 ATCAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGGACTTCGCGGAGCTGAACAAG 918
Db 790 CCGAGGCGCCCAAGCCCAAGCACTGGGCAAGCGGAGGTGACGTACCAAGCGCGCAAG 849
Qy 919 CGCACCCAGGACTTCTGGGAGTGCAGCTGGGATCCCCACCCCGCC 966
Db 850 AACCGCTGCCCGCAGAGAGGTCCGCTGGGCTACCCCGACGCGCCCG 897
```

RESULT 4
BM320900
LOCUS
DEFINITION
rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
BM320900
BM320900.1 GI:18055306
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
Mastigamoeba balamuthi.
Mastigamoeba balamuthi.
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
1 (bases 1 to 1165)
Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Senses, C. W.,
Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and
Philippe, H.

TITLE
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461

COMMENT
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 1165 Std Error: 0.00
POLYA-No. Location/Qualifiers

FEATURES
source
1..1165

```
/organism="Mastigamoeba balamuthi"  
/strain="ATCC 30984"  
/db_xref="taxon:108607"  
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"  
BASE COUNT 221 a 426 c 336 g 178 t 4 others  
ORIGIN  
Query Match 3.2%; Score 79.6; DB 13; Length 1165;  
Best Local Similarity 45.1%; Pred. No. 0.001;  
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;  
Qy 90 CATCAAGTGTTCACATGCGGCAAGGAGGCGGCACATCGCGCAACTCGCGCGCCCGC 149  
Db 46 CGTCAAGAAACAGCGTACTTCAAGCGCTTCCAGACCCAGTTCCGTGCGCGCGCGAGG 105  
Qy 150 CAAGAAGGGTGTGTGAAGTTCGGCAAGGAGGCGCCACAGATGAAGGACTGCACCGAGCG 209  
Db 106 CAAGACGACTACCGCGCGCCANCTGGTGATCCAGGACAAAGAACTACACACGCC 165  
Qy 210 CAGAGCCCACTTCTTCGCGAGGACCTGGGCTTCCCGCAGGCAAGCGCGCGAGTTCC 269  
Db 166 CAAGTACCGCTTCTGTC-----GTCCGCTTCAACAACAGGAGACATCTCTCCAGATCG 219  
Qy 270 CAGGAGCAGAACCGCGCAACAGCCCGCACAGCGCGAGCTGCAGGTGCGCGCGACAA 329  
Db 220 CTAGCCCAAGATCGACGGCGACCATCTCTGCGCGCGCTACTTCGACAGAGTCAACCG 279  
Qy 330 CCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 389  
Db 280 CTTGCGGCTCAAGTCTCGGCTGACCAACTACGCGCGCGCGCGCGCGCGCGCGCGCT 339  
Qy 390 GTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 449  
Db 340 GCGCGCGCGTGTCTGAAGAAGCTCAACCTCTACTTCCAAAGTACGAGGCTGTCAAGAAG 399  
Qy 450 CACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 509  
Db 400 CAAGCGGAGGACTACAACTCGAGGAGCTCGAGGAGCGCGCGCGCGCGCGCGCGCG 457  
Qy 510 GATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 569  
Db 458 -TGCTGACGTGGCGCTGGTGGCGACCTGAGTGGCGCGCGCGCGCGCGCGCGCGCG 515  
Qy 570 GATCTCGCGCAAGAAGCGCATCGGCACCGTGTGATCGCGCGCGCGCGCGCGCGCG 629  
Db 516 GGGCATGTGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 575  
Qy 630 CGGCGCGCAACATGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 689  
Db 576 CGCGGACAGAAGGAGTCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 635  
Qy 690 GACCGTGGCGCGTGAAGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 749  
Db 636 CGCGCGGTACATGAAGCTCTCAAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCG 695  
Qy 750 GACCGGAGGAGAGATCAAGCGCGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCG 809  
Db 696 GCGCTACGCGCAAGGAGGTGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 755  
Qy 810 GATCACCAAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 869  
Db 756 -----CAAGCAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 809  
Qy 870 GGACAGCAGCAAGTGGCGCGCAAGCTGTGGACTTTCGCGGAGCTGAACAGCGCGCG 929  
Db 810 GGGCGCAAGCCCAAGCACTGGGGCAAGCGCAGGCTGACGTACCAAGCGCGCGCG 869  
Qy 930 CTTCTGGGAGTTCAGTGGCGATCCCCACCGCGCGCGCGCGCGCGCGCGCGCGCG 989  
Db 870 CGTGGCGCGCAAGAGGTTCGCGCTGGGCTACCGCGCGCGCGCGCGCGCGCGCG 929  
Qy 990 GACCGTGTGAGGTGGCGGAGCGCTACTTACGCGTGGCGCGCGCGCGCGCGCG 1039
```


ORIGIN		Query Match 3.0%; Score 75; DB 13; Length 862; Best Local Similarity 46.2%; Pred. No. 0.0057; Matches 283; Conservative 0; Mismatches 327; Indels 3; Gaps 1;	
QY	1728	CGCGAAGATGCGCAGCGCCACCAACAGAGCTGAAGACGTGACCGAGCGGTGCAGAA	1787
DB	220	CGCGAAGATGCGGTGCTGCTGCAACATCTTCTCGACGAGGACCAACGCGCGCGC	279
QY	1788	GATCGCATGGAGAGCATCGTCTGGGGCAAGACCCCAAGTTCGCCCTGCCCATCCA	1847
DB	280	CATTCGCGAGCGGAGTCTCGGTCTTTCGCCGGAAGGCGGAGAACCTCCAGAGTACTG	339
QY	1848	GAGGAGACCTGGGAGACCTGCTGGACCGACTACTTGGCAGGCGACCTGGATCCCGAGTG	1907
DB	340	GGAGTCACTNGAAGGCCCTGTCTTCGCCCTTACCAGGACCTCAGANCACTCGTCGA	399
QY	1908	GGAGTTCTGTAACACCCCCCTCGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCAT	1967
DB	400	CGACGCGGTGACGCGACTCTAATGATCCACAAGGGTTTCGCGCGGAGGACAACCCCA	459
QY	1968	CATCGCGCGGAGACCTTCTACGTGAGCGGCGCGCCCAACCGAGACCAAGATCGGCA	2027
DB	460	GCTGCTGGAGAGCAGGAGGGCTCGAGAGGTGCGCTTCAACAACGTGCTCAAGCA	519
QY	2028	GGCGGCTACGTGACCGACCGCGGCGCGCAGAAAGATCGTGAAGCTTGACCGAGACCA	2087
DB	520	GGTCCAGAAGGAGCAGCCCGGCTTCTGGCACAAGATCTCCCGAGATCCGCGGTGAC	579
QY	2088	CCAGAAGACCGAGCTGCAGGCCATCCAGCTGGCCCTTCGAGCAGACGCGCAGGTGAA	2147
DB	580	CGAGGAGACGACACTTGGCGTGTATGAGGCTGTACCAAGCTGACCCGACGCAAGCTGT	639
QY	2148	CATCGTGACCGACGACCTAGTACGCTTGGGATC--ATCCAGGCGCGCGGCAAGAG	2204
DB	640	GTTCGCGCGCTCAACGCTCAAGACTTGTACCAAGAGCAAGTTGTAACAATCTACGG	699
QY	2205	CGAGAGCGAGCTGGTGAACAGCATCATCGACGCTGTATCAAGAGGAGAGGTGTACCT	2264
DB	700	CTCCGCCACTCGCTCATCGCGCATCAAGCGCGCACCAGCTGTGCTCGCGGCA	759
QY	2265	GAGCTGGGTGCCGCCACAGGGCATCGCGCGCAACGAGCAGATFCACAGCTGGTGAG	2324
DB	760	GPTCGCGCTGCTCGCGGCTACGCGGACGTGGGCAAGGCTGCGCGGAGTCTGCTCGCGG	819
QY	2325	CAAGGCGATCCGC 2337	
DB	820	CCAGGCTGCCGC 832	
RESULT 8		BM321393 853 bp mRNA linear EST 03-JAN-2002	
LOCUS		rockefeller.0.1222 Mastigamoeba balamuthi lambda ZAP II Library	
DEFINITION		Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA sequence.	
ACCESSION		BM321393	
VERSION		BM321393.1	
KEYWORDS		EST.	
SOURCE		Mastigamoeba balamuthi.	
ORGANISM		Mastigamoeba balamuthi	
REFERENCE		1 (bases 1 to 853)	
AUTHORS		Bapteste, E., Brinkmann, H., Lee, J. A., Moore, D. V., Sensen, C. W., Gordon, P., Durufle, L., Gaasterland, T., Lopez, P., Muller, M. and Philippe, H.	
TITLE		The analysis of 100 genes supports the grouping of three highly divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba	
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)	
MEDLINE		21819461	
COMMENT		Contact: Muller Miklos Laboratory of Biochemical Parasitology	
FEATURES		source	
BASE COUNT		175 a 294 c 260 g 126 t 7 others	

The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockefeller.edu
Insert Length: 853 Std Error: 0.00
POLYA=No.

FEATURES

Location/Qualifiers
1..853
/organism="Mastigamoeba balamuthi"
/strain="ATCC 30984"
/db_xref="taxon:108607"
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
/note="syn: Phreatamoeba balamuthi"

BASE COUNT 156 a 307 c 263 g 125 t 2 others

Query Match 3.0%; Score 73.2; DB 13; Length 853;

Best Local Similarity 46.9%; Pred. No. 0.011;

Matches 299; Conservative 0; Mismatches 329; Indels 9; Gaps 2;

Qy 388 CTGTGGGAGCGCCCTGGTGAGCATCAAGGTGGCGGCCAGATCAAGGAGGCCCTGTGCTG 447
Db 125 CTGGCCCGCGCTGTCTGAAGAAGCTCAACCTCGACTCCAAGTACGAGGGTGTCAAGAAG 184
Qy 448 GACACCGCGCGGACGACACCGCTGTCTGAGAGATGAGCTGCCCGGCAAGTGGAGGCC 507
Db 185 GTCAACGCGGAGGACTACAACGTGAGAGCTCGACGACGGCGCCGCCCTTCAAGGCC 244
Qy 508 AAGATGATCGCGCGCATCGCGCGCTTCAATCAAGGTGGCGGCTAGTACGACAGATCTGTATC 567
Db 245 CTGCTCGAGCTCGGCTGTCTGCGACCTGACGCGCGCGCGGCTTTCGCGCCCTCAAG 304
Qy 568 GAGATCTCGGCAAGAAGGCCATCGGACCGTGTGTGATCGGCCCGCCCGCGTGAACATC 627
Db 305 GGCATGTGCGAGCGCGCGGTCAACGTCCCG---CACACGAGACCGCGTTCGCGCTTC 361
Qy 628 ATCGGCCCAACATGCTGACCCAGCTGGCTGGCTGACCCCTGACCTTCCCATCAGCCCCATC 687
Db 362 AACGGCGCAAGAAGGAGCTCAACGCGCGCTTCTCGCAAGTATCTTTCGCGGCCAC 421
Qy 688 GAGACCTGTCGCGGTGAAGCTGAAGCGCGCATGGAAGCGCGCCCAAGGTGAAGCAGTGGCC 747
Db 422 GTCGCGGCTACATGAGCTCTCAAGGAGGAGCGCGCGCGCTTCGACCGCGCATTC 481
Qy 748 CTGACCGAGGAGAGATCAAGCGCCTGACCGCCATCTCGAGGAGATGGAGAAGGAGGC 807
Db 482 TCGCGCTACGCAAGAGAGGTGTACCGCGCGACATGCTCGAAGATCTACACCGAGGCC 541
Qy 808 AAGATCACCAGATCGGCGCGCGAGACCCCTTACACACCGCGCGTGTTCGCCATCAAGAAG 867
Db 542 CA-----CAAGCAGATCGCGCGCGCGACCTTCGTCGCCCAAGCGCGCTCGAAGGCC 595
Qy 868 AAGCAGACGACCAAGTGGCGCAAGCTGGTGACTTCCGCGAGCTGAACAGCGCACCCAG 927
Db 596 GAGGCGCCCAAGCCCAAGCACTGGGCAAGCGCAGGCTGACGTACCAGCGCGCAAGAC 655
Qy 928 GACTTCTGGAGGTGACGTGGGATCCCCAACCGCGCGCTGAAGAAGAAGAGAGC 987
Db 656 CGCGTCCCGCAGAGAAGTCCGCTGGGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCG 715
Qy 988 GTGACCGTGTGGAGTGGGCGAGCGCTTTCAGG 1024
Db 716 CCGGTCGCTGTACACNGCCATCTCCGCTGCGGCG 752

RESULT 9
BM587428/c

LOCUS
DEFINITION

1700687321202 A.Gam.ad.cdna..blood1 Anopheles gambiae cDNA clone
19600449696110 5', mRNA sequence.

ACCESSION
VERSION

BM587428
BM587428.1 GI:18883289

KEYWORDS
SOURCE

EST.
African malaria mosquito.

ORGANISM

Anopheles gambiae

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
Anopheles.

REFERENCE

AUTHORS

1 (bases 1 to 566)
Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
Celera Anopheles gambiae EST project
Unpublished (2002)

TITLE

JOURNAL

COMMENT

Contact: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004AX row: G column: 08
Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
1..566

source

/organism="Anopheles gambiae"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"

/db_xref="taxon:7165"

/clone="19600449696110"

/clone_lib="A.Gam.ad.cdna.blood1"

/dev_stage="Adult"

/lab_host="DH10b"

/note="Vector: pSport1; site_1: SalI; site_2: NotI; Whole
adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"

BASE COUNT 60 a 202 c 165 g 139 t

ORIGIN

Query Match 3.0%; Score 72.8; DB 13; Length 566;
Best Local Similarity 48.3%; Pred. No. 0.013;
Matches 203; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

Qy 1950 GCTGGAGAGAGGCCATCATCGGCGCGGAGACCTTCTACGTGGACGGCGCGCAACCG 2009
Db 482 GTTCGAGATCGACGCCAACGGCATCTGCAGGTGTGCGCGGAGGACAGGGCACGGGCA 423
Qy 2010 CGAGACCAAGATCGCAAGGGCGGTAGTGACCGACCGGGCGCGGAGAGATCGTAG 2069
Db 422 CCGGAAAGATCGTCTACCAACGACGACGACGACGCGCTGACGCGGAGCATCGAGCG 363
Qy 2070 CCTGACCGAGACCAACACGAGAGCGAGCTGCGAGGCCATCCAGCTGGCCCTCAGGA 2129
Db 362 CATGATCAAGATCGGAGCGGTTCCGCGCGAGCGACGACAGAGCTGAAGGAGCGGTGA 303
Qy 2130 CAGCGGACGAGGTGAACATCGTGACCGACGAGCAGTACGCCCTGGGCATCATCCAGC 2189
Db 302 GGCCCGCAACGAGCTCGAGAGCTACGCGCTACAGCTGAAGAACCAGCTCAGCTCAAGA 243
Qy 2190 CCAGCCCGACAGAGCGAGAGCGAGCTGGTGAACGAGATCATCGAGAGCTGATCAAGAA 2249
Db 242 CAAGCTGGCGGAGAGCTGTCGACGACGACGACGACGACGACGACGACGACGACGACG 183
Qy 2250 GGAGAAGTGTACCTGAGCTGGTGGCGCGCGCAAGGCGCATCGCGCGCAACGAGCAGAT 2309
Db 182 GAAGATCAAGTGGCTGGACGAGAGACGAGGACCGAGGCGGAGAGTACAAGACGAGAA 123
Qy 2310 CGACAAGCTGTGAGCAGGAGCATCCGCAAGTGTGTCTTGGAGCGCATCGATGGCGG 2369
Db 122 GAAGGAGCTGGAGAGATCGTGCAGGCCCATCATTTGCCAAGCTGTACGCGAGCAGTGGCGG 63

RESULT 10

BE601575

LOCUS

DEFINITION

BE601575 640 bp mRNA linear EST 22-OCT-2001
HVSMEH0098K08f Hordeum vulgare 5-45 DAP spike EST library

HVCDNA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEh0098K08f, mRNA sequence.

ACCESSION BE601575
VERSION BE601575.3 GI:16322423
KEYWORDS EST.

REFERENCE	SOURCE	ORGANISM
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

mordeum vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.

REFERENCE

Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton, R.D., Close, S.J., Gates, R. and Main, D.

Development of a genetically and physically anchored EST resource for barley genomics: morex 5-45 DAP spike cDNA library

JOURNAL Unpublished (2001)

On Aug 21, 2000 this sequence replaced gi:13190104.

COMMENT
On Aug 21, 2000
Contact: Wing PA

CONTACT: WING RA
Clemson University, Clemson, SC 29634-0901

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, C

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total ha bases = 461

total 119 bases = 401
 seq primer: ATTATCCCTCCTAATAACCG

seq primer: AATTAACTCTCACTAAAGGG
right amplification primer: 524

FEATURES
source

```
1. 1040
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEH0098K08f"
/clonelib="Hordeum vulgare 5-45 DAP spike EST library
HVCNAN009 (5 to 45 DAP)"
```

```

nvcumv003 (J CO 4J bar)
/tissue_type="5-45 DAP Spike"
/lab_host="SOLR"
/note="vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
plants were grown in the greenhouse at the University of

```

Plants were grown in the greenhouse at the University of California, Riverside (Fenton, SJ Close, TJ Close). Whole spikes with awns trimmed were collected at 5, 10, 15, 20, 30 and 45 DAP (Fenton). Total RNA was prepared from each

pool, equal quantities of all six RNA pools were combined, 30 and 45 DAF (Fencon). Total RNA was prepared from each pool, equal quantities of all six RNA pools were combined, poly(A)⁺ RNA was purified from the mixture, one primary

poly(A) RNA was purified from the mixture, and one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids

(Choi) in the TJ Close lab at the University of California, Riverside. Phagemids were plated and picked at the Clemson

University Genomics Institute (UGI) (Begum, Palmer, Frisch, Atkins and Wing) plasmid DNA preparations, DNA

sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).

849	Qy	CGTGTTCGCCATCAGAAGAAGGACAGCACCAGTGGCGCAAGCTTGGTCTCGCGA	908
170	Db	CTGAGCCGCAACTTCAAGCACTCAACCTTGACCTTCCAGCTGCAGGACGGCGGGCGCA	229
909	Qy	GCTGAACAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCGATCCCCACCCCGCCGG	968
230	Db	GCTCAAGGTGGACGCTGTTCGGCACCCCGCGCACCATTGGCGGCATCCGCACCGCCAT	289
969	Qy	CTGTAAAGAAGAAGAGCGGTGACCGTGTGGACGTGGCGAGCGCTACTTCAGCGTGCC	1028
290	Db	CTCCACGCTCCAGAACCTCATCACCGCGCTACCAAGGCGTTCGCTACAAGATGCGCTT	349
1029	Qy	CTGTGACGAGGACTTCCGGAAGTACACCGCTTCACCATTCCCGAGCATCAACACGAGAC	1088
350	Db	CGTCTACGGTCACTTCCCCATCAACGCTTCCATCACGCGCGCAAGCGGGGCTCAGAT	409
1089	Qy	CCCGGCACTCCGCTACCACTACAAGCTGTGCCCCAGGCGTGAAGGGGACGCCACGAT	1148
410	Db	CCGCAACTTCTTCGGCGAAGAAGGTGAGGAAGGTGGACATGCTCGACGGGTCACCAT	469
1149	Qy	CTTCCAGACGAGCATGACCAAGATCTGGAGCCCTTTCGCGGCCGCAACCCCGAGATCGT	1208
470	Db	CTTGGCGTCCGACAAGTCAAGATGAGATCGTCTTCGACGGGCAACGACATCGAGTCGT	529
1209	Qy	GATCTACCAAGGCCCCCT	1226
530	Db	NTTCGCTCCGGCGCCT	547

	BM372120	EST	23-JUL-2002
RESULT 11	LOCUS	linear	
	BM372120	mRNA	
	DEFINITION	500 bp	
		EBR03_SQ004.B09.R root.	
		waterlogged, cv Optic, EBR03	
		Hordeum vulgare cDNA clone EBR03_SQ004_B09_5'	
		, mRNA sequence.	
	ACCESSION		
	BM372120		
	VERSION		
	BM372120.2		
	GI:21948499		
	KEYWORDS		
	EST		

ESI.
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
1 (bases 1 to 500)
Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,

TITLE
Ramsay, L., Machray, G., Marshall, D.F.M. and Waugh, R.
Development of Barley Transcriptome Resources

JOURNAL Unpublished (2001)
COMMENT On Jan 10, 2002 this sequence version replaced gi:18115510.

Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology

Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
E-mail: 08 44 1200 55333

Tel: 00 44 1382 562731
Fax: 00 44 1382 562426

Email: est@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse

```

BASE COUNT      108 a 177 c 139 g 76 t
ORIGIN

Query Match
Best Local Similarity 2.9%; Score 71.6; DB 13; Length 500;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGGAGGAGGCAAGATCACCAGATCGCGCCCGGAGAACCCCTACAAACACCCC 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 GGAGAGGTGACGGTGAGAGGTGTCGGCCAAAGATGATCTCGGTGACGGGGCCCGCGGCAC 91
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 849 CGTGTTCGCCATCAAGAAGAAGGACAGCACCAGTGGCGCAAGCTGGTGAGCTTCGGCA 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 CTTGACCCGCAACTTCAAGCACTCAACCTCGACTTCAGCTGACAGTGCAGGACGGCGGCA 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 909 GCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGATCCCGCCACCGCGGG 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 GCTCAAGGTGACGCTGCTGGGACCCGCGCCACCATGCGCGGCATCCCGACCGGCAT 211
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 969 CTTGAAGAAGAAGAGCGGTGACCGTCTGGACGTGGCGCGACCGCTACTTCAGGCTGCC 1028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 CTCCACCTCAGAACCTCATACCGGGCTCACAAGGGCTTCGCTACAAGATCGGCTT 271
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 CTTGACAGGACTTCCCAAGTACACCGGCTTCACCATCCCGAGCATCAACAACGAGAC 1088
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 CGTCTACGCTCACTTCCCGCATCAAGCGCTCCATCACCSCCGCCCAACCGGGCATCGAGAT 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1089 CCCCAGGATCGCTACCAAGTACAGGTCTGCCCGGAGGTGGAGGCGACGCCCGAGCAT 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 CCGCAACTTCTCGCGGAGAGAGGTGAGGAGGTGGACATGCTCGACGGGTCACCAT 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 CTTCCAGAGCATGACCAAGATCCTGGAGCCCTTCGCGCCGCGCAACCCGAGATCGT 1208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 CTTGCGGTCCGAGAGGTCAGAGTATGATCGTCTCTGACGGCAACGACATCGAGCTCGT 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1209 GATCTACAGGCCGCCCT 1226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 CTCCCGTCCGCGGCCCT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
BM368580 538 bp mRNA linear EST 23-JUL-2002
LOCUS EBem08_SQ004_A06_R embryo, 40 DPA, no treatment, cv Optic, EBem08
DEFINITION Hordeum vulgare cDNA clone EBem08_SQ004_A06 5', mRNA sequence.
ACCESSION BM368580
VERSION
KEYWORDS EST.
SOURCE BM368580.2 GI:21937722
ORGANISM Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 538).
Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
Development of Barley Transcriptome Resources
Unpublished (2001)
On Jan 10, 2002 this sequence version replaced gi:18111970.
Contact: Waugh R, Marshall DF
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers
1..538
/organism="Hordeum vulgare"
/cultivar="Optic"

FEATURES
source

```

```

/db_xref="taxon:4513"
/clone="EBem08_SQ004_A06"
/clone_lib="embryo, 40 DPA, no treatment, cv Optic,
EBem08"
/tissue_type="embryo"
/dev_stage="40 DPA"
/lab_host="DH108"
/Note="Vector: pSPORT1; Site 1: Sal I; Site 2: Not I;
Non-normalised library, directionally cloned into pSPORT1.
Derived from embryos dissected from developing grains (40
days post anthesis) in glasshouse grown barley plants.
Developed as part of the barley transcriptome resources of
BBSRC/SEERAD funded cereal IGF (Investigating Gene
Function) project."
BASE COUNT      107 a 205 c 149 g 77 t
ORIGIN

```

```

Query Match
Best Local Similarity 2.9%; Score 71.6; DB 13; Length 538;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAGGAGGAGGCAAGATCACCAGATCGCGCCCGGAGAACCCCTACAAACACCCC 848
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93 GGAGAGGTGACGGTGAGAGGTGTCGGCCAAAGATGATCTCGGTGACGGGGCCCGCGGCAC 152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 849 CGTGTTCGCCATCAAGAAGAAGGACAGCACCAGTGGCGCAAGCTGGTGAGCTTCGGCA 908
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 153 CTTGACCCGCAACTTCAAGCACTCAACCTCGACTTCAGCTGACAGTGCAGGACGGCGGCA 212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 909 GCTGAACAGCGCACCCAGGACTTCTGGGAGGTGAGCTGGGATCCCGCCACCGCGGG 968
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 213 GCTCAAGGTGACGCTGCTGGGACCCGCGCCACCATGCGCGGCATCCCGACCGGCAT 272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 969 CTTGAAGAAGAAGAGCGGTGACCGTCTGGAGCCCTTCGCGCCGCGCAACCCGAGATCGT 1028
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 273 CTCCCGTCCGAGAGGTCAGAGTATGATCGTCTCTGACGGCAACGACATCGAGCTCGT 332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 CTTGACAGGACTTCCCAAGTACACCGGCTTCACCATCCCGAGCATCAACAACGAGAC 1088
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 333 CGTCTACGCTCACTTCCCGCATCAAGCGCTCCATCACCSCCGCCCAACCGGGCATCGAGAT 392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1089 CCCCAGGATCGCTACCAAGTACACCGGCTTCGCGCCGCGCAACCCGAGATCGT 1148
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 393 CCGCAACTTCTCGCGGAGAGAGGTGAGGAGGTGGACATGCTCGACGGGTCACCAT 452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 CTTGACAGGAGCATGACCAAGATCCTGGAGCCCTTCGCGCCCGCCCAACCCCGAGATCGT 1208
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 453 CTTGCGGTCCGAGAGGTCAGAGTATGATCGTCTCTGACGGCAACGACATCGAGCTCGT 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1209 GATCTACAGGCCGCCCT 1226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 513 CTCCCGTCCGCGGCCCT 530
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 13
AJ471121 540 bp mRNA linear EST 24-MAY-2002
LOCUS AJ471121 S00008 Hordeum vulgare cDNA clone S0000800191A10F1, mRNA
DEFINITION sequence.
ACCESSION AJ471121
VERSION AJ471121.1 GI:21187077
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 540).
Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
Barley EST's
Unpublished (2002)
Contact: Schulman AH
Institute of Biotechnology

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:14:58 ; Search time 80 Seconds
(without alignments)
9441.811 Million cell updates/sec

Title: US-09-610-313-31
Perfect score: 2463
Sequence: 1 gtcgacgcccattggccga.....gggtagcaccggtgaattc 2463

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.8	44.9	2601	4	US-09-117-217-7
2	1106.8	44.9	2601	4	US-09-117-217-9
3	1106.8	44.9	2601	4	US-09-117-217-11
4	1106.8	44.9	2601	4	US-09-117-217-13
5	1100.4	44.7	7399	2	US-08-418-848A-9
6	1100.4	44.7	9709	2	US-08-188-583-5
7	1100.4	44.7	9709	3	US-08-388-353-1
8	1100.4	44.7	9709	3	US-08-488-551B-1
9	1100.4	44.7	9709	4	US-09-309-572-15
10	1100.4	44.7	12494	4	US-08-935-312-13
11	1100.4	44.7	12494	4	US-08-848-760B-33
12	1100.4	44.7	15581	3	US-08-646-538-35
13	1100.4	44.7	15581	4	US-09-503-222-35
14	1090.8	44.3	9737	2	US-08-944-449-7
15	1090.8	44.3	9737	4	US-09-353-362-7
16	1055.6	42.9	9746	1	US-08-022-835-3
17	1055.6	42.9	9746	1	US-08-388-809-3
18	1055.6	42.9	9746	2	US-08-647-714-3
19	1054.4	42.8	8932	4	US-09-124-900-1
20	1054.4	42.8	8933	3	US-08-463-210-4
21	1054.4	42.8	8933	4	US-09-620-958A-3
22	1054.4	42.8	8933	4	US-09-620-958A-4
23	1054.4	42.8	8933	4	US-09-620-958A-9
24	1042	42.3	9739	1	US-08-022-835-1
25	1042	42.3	9739	1	US-08-388-809-1
26	1042	42.3	9739	2	US-08-647-714-1
27	1036.8	42.1	5362	3	US-08-463-210-5

28 1025 41.6 3033 2 US-07-743-357-19 Sequence 19, Appl
29 1025 41.6 4113 2 US-07-743-357-21 Sequence 21, Appl
30 1019.2 41.4 9207 3 US-08-388-353-800 Sequence 800, App
31 1019.2 41.4 9207 3 US-08-488-551B-800 Sequence 800, App
32 929.4 37.7 3856 2 US-07-743-357-20 Sequence 20, Appl
33 924 37.5 2739 2 US-07-743-357-16 Sequence 16, Appl
34 843.6 34.3 2348 2 US-08-876-546A-1 Sequence 1, Appl
35 843.6 34.3 2348 4 US-09-412-252-1 Sequence 1, Appl
36 843.6 34.3 2348 1 US-09-079-675-1 Sequence 1, Appl
37 831 33.7 9793 1 US-08-470-202-56 Sequence 56, Appl
38 831 33.7 9793 1 US-08-471-770-56 Sequence 56, Appl
39 831 33.7 9793 2 US-08-468-059-56 Sequence 56, Appl
40 831 33.7 9793 4 US-09-109-916-56 Sequence 56, Appl
41 823.4 33.4 1680 5 PCT-US93-06748-1 Sequence 1, Appl
42 604 24.5 1256 4 US-09-158-695-18 Sequence 18, Appl
43 553 22.5 3168 2 US-08-659-251-14 Sequence 14, Appl
44 553 22.5 3168 4 US-09-256-490-14 Sequence 14, Appl
45 553 22.5 3168 5 PCT-US96-11445-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-09-117-217-7

; Sequence 7, Application US/09117217

; Patent No. 6221578

; GENERAL INFORMATION:

; APPLICANT: de BETHUNE, Marie-Pierre

; APPLICANT: HERTOGS, Kurt

; APPLICANT: PAUWELS, Rudi

; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY

; FILE REFERENCE: 1377-125P

; CURRENT APPLICATION NUMBER: US/09/117.217

; CURRENT FILING DATE: 1998-07-24

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 2601

; TYPE: DNA

; ORGANISM: HIV-HXB2

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(492)

; OTHER INFORMATION: gag Polyprotein

US-09-117-217-7

Query Match 44.9%; Score 1106.8; DB 4; Length 2601;
Best Local Similarity 68.2%; Pred. No. 1.4e-173;
Matches 1618; Conservative 0; Mismatches 732; Indels 22; Gaps 5;

Qy 14 TGGCCGAGGCATGAGCCAGCCACCA---GCGCCAACTCTGTATGAGGCGCAGCAACT 70

Db 77 TGGTGAAGCAATGAGCCAAATTAACAAATTCAGCTACCAATATGATGAGAGGCAATT 136

Qy 71 TCAGGGCCCCAAGCGCATCATGCTTCACTGCGGCAAGGAGGCGCCACATCGCCC 130

Db 137 TTAGAACCAAGAAAGATTGTTAGTGTTCATTTGCGCAAGAGGGGCACACAGCCA 196

Qy 131 GCAACTGCGCGCCCGCCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGCGCCACACAGA 190

Db 197 GAATTGCGAGGCCCTAGGAAAGAGGCTGTGGAATGTGGAAGAGGACACACAAA 256

Qy 191 TGAAGACTGCACGAGCGCCAGCCAACTTCTTCGCGAGGAGCTGCGCTTCCCCCAGG 250

Db 257 TGAAGATTGTTACTGAGACAGACAGCTAA-TTTTTTAGGGAAGATCTGCGCTTCTTACAAG 315

Qy 251 GCAGGGCCCGAGTTCCTCCAGCAGCAGACCGCCCAACAGCCCCCAGCCCGCAGC 310

Db 316 GGAAGGCGCGGAATTTTCTTCAGAGCAGACAGCCAGCCAGCCCGCCAGAGAGAGC 375

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATON NUMBER: US/08/418,848A
FILING DATE: 07-APR-1995
CLASSIFICATION: 526
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/936,849
FILING DATE: 28-AUG-1992
CLASSIFICATION: 526
ATTORNEY/AGENT INFORMATION:
NAME: RICHARD W. BORK
REGISTRATION NUMBER: 36,459
REFERENCE/DOCKET NUMBER: 2026-40910S2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 7399 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-418-848A-9

Query Match 44.7%; Score 1100.4; DB 2; Length 7399;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;
QY 14 TGGCCGAGGCGATGAGCCAGGCGACCA--CGGCCACATCTTGATGAGGCGAGCAACT 70
DB 1838 TGGCTGAAGCAATGAGCAAGTACCAATCCAGTACATATGATACAGAAAGGCAATT 1897
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTCAACTCGCGCAAGGAGGCGCACATCGCC 130
DB 1898 TTAGGAACCAAGAAAGACATGTTAAGTGTTCATTTGCGCAAGAGGCGACATGCCA 1957
QY 131 GCAACTCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGCAGAGGAGGCGCCACAGA 190
DB 1958 AAAATTGCGAGGCGCCCTAGGAAAGAGGCTGTGGAATGTGGAAGAGGAGGACCAAA 2017
QY 191 TGAAGGACTGACCGAGCGCCAGGCAACTTCTTCGCGAGGAGCTGGCCCTTCCCGCAGG 250
DB 2018 TGAAGGATTGTACTGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGGCCCTTCCCAAG 2076
QY 251 GCAAGGCGCGAGTTCTCCCGAGGAGCAGAAACCGCGCAACAGCCCGCAGCGCGAGC 310
DB 2077 GGAAGGCGAGGAATTTCTTCAGAGCAGACAGGAGCCAGCGCCACAGAGAGAGC 2136
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGAGCGCGCAGGCA 364
DB 2137 TTCAGGTTTGGGGAAGAGAGCAACAACCTCCCTCTCAGAAGAGGAGGCGATAGAGAGAA 2196
QY 365 -----CCCTCAACTTCCCGCAGATACCTGTGGAGCGCCCGCTGTGAGCATCAAGG 418
DB 2197 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCTGTCACAATAAGA 2256
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCAGCAGACCGCTGTGGAGG 478
DB 2257 TAGGGGCGCAATTAAGAGAGCTCTATTAGATACAGAGCAGATGATACAGATTAAGAG 2316
QY 479 AGATGAGCGCTCGCGCGCAAGTGAAGCCCAAGATGATGCGGCGGATCGCGGCGCTTCATCA 538
DB 2317 AAATGAATTTGCCAGGAGATGGAAGCAAAAATGATAGGGGGAATTTGGAGTTTATCA 2376
QY 539 AGGTGCGCGAGTAGACACAGATCTGTATCGAGATCTCGGCAAGAGGCGCATCGGACCG 598
DB 2377 AAGTAGGACAGTAGATCAGATCACTATAGAAATCTCGGACATCAAGCTATAGGTACAG 2436

QY 599 TGCTGATCGGCGCCACCCCGTGAACATCATCGGCCCAACATGCTGACCCAGCTGGCT 658
DB 2437 TATTAGTAGGACCTACACCTGTCAACATATTGGAAGAAATCTGTTGACACGATTGGCT 2496
QY 659 GCACCCCTGAATCTCCCATCAGCCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCA 718
DB 2497 GCACCTTTAAATTTTCCCATAGTCTTATGAGACTGTACCAAGTAAATTAAGACCGAGAA 2556
QY 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCGCTGACCG 778
DB 2557 TGGATGCGCCCAAGTTAAACATGGCCATTGACAGAGAAATAAAGCATTAGTAG 2616
QY 779 CCATCTCGAGGAGATGGAGAAGGAGGCAAGATCAGCAAGATCGGCCCGCCGAGAACCCCT 838
DB 2617 AAATTTGTACAGAAATGGAAGGAGGAAATTTCAAAATTTGGGCTGAAATCCAT 2676
QY 839 ACAACACCCCGTGTTCGCATCAAGAAAGAGACAGCAAGTGGCGGCAAGTGGTGG 898
DB 2677 ACAATACTCCAGTATTTGCCATAAAGAAAGACAGTACTTAAATGGAGAAATTAGTAG 2736
QY 899 ACTTCCGCGAGCTCAACAAAGCGCACCCAGGACTTCTGGAGGTGACGTGGCGATCCCC 958
DB 2737 ATTTGAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAC 2796
QY 959 ACCCGCGCGCTGAAGAAGAAAGAGCGTGAACGCTGCGAGCTGGCGGACGCTACT 1018
DB 2797 ATCTGCGAGGTTAAACACAGAAATAATCAGTAACAGTACTGATGTGGCGCATGCATATT 2856
QY 1019 TCAGCTGCGCGCTGGAGGAGGACTTCCGCAAGTACACCGCTTCAACATCCCCAGCATCA 1078
DB 2857 TTTGAGTTCCTTATAGATAAAGACTTCAGGAAGTATCTGCACTTTTACCATACTAGTATAA 2916
QY 1079 ACAACGAGACCCCGGCGCTACCAAGTACAGCTGCTCCCGCAGGCTGGAGGCA 1138
DB 2917 ACAATGAGACACCGAGGATTAGATATCAGTACATCTCTCCACAGGATGGGAAAGAT 2976
QY 1139 GCCCGACGATCTCCAGAGCAGCATGACCAAGATCTTCCGAGCCCTTCCCGCCCGCAACC 1198
DB 2977 CACGAGCAATATCCAGTGTAGCATGACAAATCTTTAGAGCCTTTAGAAGAAACAAATC 3036
QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTAGCTGGGAGCGACCTCGAGATCG 1252
DB 3037 CAGACATAGTCACTCAATACATGATGATGATGATGATGATGATGATGATGATGATGATG 3096
QY 1253 GCCAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGCTTCACCA 1312
DB 3097 GGCAGATAGAACAAATAGAGGAACCTGAGACACATCTCTTTGAGTGGGATTTTACCA 3156
QY 1313 CCCCCGACAGAGCAGCAGAGGAGGCGCCCTTCCCTGTGATGGCTTACGAGCTGCACC 1372
DB 3157 CACGAGCAAAACATCAGAAAGAACCTCCATTTCTTGGATGGTTTATGACTCCATC 3216
QY 1373 CCGACAGTGAACCGTGGAGCCCATCGAGCTGCGCGAGAGAGAGAGCTGGACCGTGAACG 1432
DB 3217 CTGATAATGAGCAGTACAGCTTATGCTGCTGCCAGAAAGAGACAGCTGAGCTGCTCAATG 3276
QY 1433 ACATCCAGAGCTGCTGGGCAAGCTGAACCTGGCGCAGCAGATCTACCCCGCATCAAGG 1492
DB 3277 ACATACAGAAATTAGTGGGAAATTAGTGGGCAAGTACAGATTTATGACGGATTAAG 3336
QY 1493 TGGCGCGAGCTGTGAAGCTGTGCGCGCGCAGGCGCTGACCGACATCTGCGCCCTGA 1552
DB 3337 TAAGGCAATTTATGTAATCTTTAGGGGAGCAACAGGACTAACAGAGTGTAGTACCACATA 3396
QY 1553 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCGCGCTGACG 1612
DB 3397 CAGAAGAGCAGAGCTAGAACTGGCAGAAACAGGGGAGATTTCTAAAGAAACCGGTACATG 3456
QY 1613 GCGTGTACTACGACCCCAAGGAGGAGCTGCTGGCGGAGATCCAGAGGAGGCGGCGACGAC 1672
DB 3457 GAGTGTATTATGACCCCATCAAGACTTAATAGCAAGAAATACAGAGGAGGCGGAGGCC 3516
QY 1673 AGTGGACCTACAGATCTACGAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1732

[illegible]

RESULT 6

```

US-08-188-583-5
; Sequence 5, Application US/08188583
; Patent No. 5851813
;
; GENERAL INFORMATION:
;
; APPLICANT: Destosiars, Ronald C.
;
; TITLE OF INVENTION: PRIMATE LENTIVIRUS VACCINES
;
; NUMBER OF SEQUENCES: 57
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188, 583
; FILING DATE:
;
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
;

```

QY 719 TGGAGGGCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACCG 778
DB 2596 TGGATGGCCCCAAAAGTTAAACAAATGGCCATTGACAGAGAAAAATAAAGCAATTAGTAG 2655
QY 779 CCACTCGGAGGAGNATGAGAGAGGAGGCAAGATCACCAGATCGGCGCCCGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGAAAGAGGAGGAGAAATTTCAAAAATTTGGGCTGAAATCCAT 2715
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
DB 2716 ACAATCTCCAGTATTTGCCATAAAGAAAAGACAGTACTAAATGGAGAAAATTAGTAG 2775
QY 899 ACTTCGGGAGCTGAACAGGCAAGCCAGGACTTCTGGGAGGTGAGCTGGGCAATCCCCC 958
DB 2776 ATTTTCAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAATTCATTAATAGGAATACCAC 2835
QY 959 ACCCGCGCGCTGAGAGAGAGAGAGAGCGTGAACGCTGCTGAGCTGGGCGACGCTACT 1018
DB 2836 ATCCTGCGAGGTTAAACAGAGAAAATCAGTAACAGTACTGATGTGGGCGATGCATATT 2895
QY 1019 TCAGCGTGCCTTGACGAGGACTTCCGCAAGTACACGCGCTTCACCAATCCCGAGCATCA 1078
DB 2896 TTTTCAGTCCCTTAGATAAAGACTTCAGGAAGTATAGTGCATTTACCAATACCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGCTACCGTACCAAGTGTGCTGCGCCAGCGCTGGAAGGCA 1138
DB 2956 ACAATGACACACAGGATTAGATATACATAGTACAATGTCTTCCACAGGATGGAAGGAT 3015
QY 1139 CCCCAGAGATCTTCAGAGAGAGATGACCAAGATCTCGGAGCCCTTCGCGCGCGCAACC 1198
DB 3016 CACCAGCAATATTCAGTGTAGTACGACAAAAATCTTAGAGCCTTTTAGAAAAACAAATC 3075
QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTAGCTGGCGAGCAGCTGGAGTCG 1252
DB 3076 CAGCATAGTCTATCTATCAATACATAGTATGATTTGTATGATAGTCTGACTTAGAAATAG 3135
QY 1253 GCCAGCACCGGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGCTTCACCA 1312
DB 3136 GGCAGCATAGAAAAAATAGAGAACTGAGCAACATCTGTTGAGGTGGGATTTACCA 3195
QY 1313 CCCCAGCAAGAACACACAGAGAGGAGCCCCCTTCCTGTGATGGCTACGAGCTGCAC 1372
DB 3196 CACCAGCAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTTATGAACCTCCATC 3255
QY 1373 CCGACAAGTGACCGCTGACGCCCTGAGCTGCGGAGAGGAGCTGGACCGTGAAG 1432
DB 3256 CTGATAATGACAGTACAGCCTTAGTGTGCGCAAGAAAGACAGCTGGAGTGTCAATG 3315
QY 1433 ACATCCAGAAAGCTGGTGGCAAGCTGAATCTGGGCGAGCCAGATCTACCCCGCATCAAG 1492
DB 3316 ACATACAGAAATAGTGGGAAATTTGAATTTGGCAAGTCAGATTTATGCGAGGATTAAG 3375
QY 1493 TGGCCAGCTGTCAAGCTGTGCGCGCGGCGCAAGGCCCTACCCAGCATGTGCCCCGTA 1552
DB 3376 TAAGCAATATGATAACTTTTAGGGGAACCAAGCAAGCACTAACAGAAAGTAGTACCCTAA 3435
QY 1553 CCGAGGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGGAGCGCGTGCACG 1612
DB 3436 CAGAAGAGCAGAGCTAGAACTGCGAGAAACAGGGAGATTTCTAAAGAACCGGTACATG 3495
QY 1613 GCGTGTACTAGACCCCAAGAACGCTGTGGCGCGAGATCCAGAAGCAGGCGCCACGACC 1672
DB 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGGCAAGGCC 3555
QY 1673 AGTGGACCTACAGATCTACAGAGGCTTCAGAGACCTCAAGACCGGCAAGTACGCCA 1732
DB 3556 AATGACATATCAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAAATATGCA 3615
QY 1733 AGATGCGCACCCCAACCAAGCAGTGAAGCAGCTGACCGAGGCGGTGAGAGATCG 1792
DB 3616 GAATGAAGGGTGGCCACACTAATGATGTGAACAATTAACAGAGCGAGTACAAAAATAG 3675
QY 1793 CCATGGAGAGCATCGTGTCTGGGCAAGACCCCAAGTTCCGCTGCCCTCCATCCAGAGG 1852

DB 3676 CCACAGAAAGCATAGTAATAATGGGAAAGACTCTCTAAATTTAAATTTACCATACAAAGG 3735
QY 1853 AGACCTGGAGAGACTGTGGACCGGACTACTTGGCAGGCGCACCTGGATCCCCGAGTGGAGT 1912
DB 3736 AAACATGGGAAGCATGGTGGACAGAGTATTTGGCAAGCCACCTGGATTTCTGAGTGGAGT 3795
QY 1913 TCGTCAACACCCCGCTGGTGAAGCTGTGTACAGCTGAGAGAGGAGCCCATCATCG 1972
DB 3796 TTGTCAATACCCCTCTTATGTAGTGAAGTTATGTACAGTTAGAGAAAGAACCCATAATAG 3855
QY 1973 GCGCGGAGACCTTCTACGTGGACGGCGCCCAACCGCGGAGACCAAGATCGGCAAGCGCG 2032
DB 3856 GAGCAGAAACTTTCTATGTAGTGGCGAGCCCAATAGGGAACCTAAATTTAGGAAGCAG 3915
QY 2033 GCTAGCTGACCGACCGGGCGGCGAGAGATCGTGAGCTGACCGAGAGACCAACACAGA 2092
DB 3916 GATATGTAACGTACAGAGAGAGACAAAAAGTTGTCCCTTAACGGACACACAACAAATCAGA 3975
QY 2093 AGACCGAGCTGCAGGCCATTCAGCTGGCCCTGCGAGSACAGCGGAGGTGAACATCG 2152
DB 3976 AGACTGAGTTTACAGCAATTCATCTAGCTTTGCAGGATTTAGAAAGTAAACATAG 4035
QY 2153 TGACCGACAGCGAGTACGCCCTTGGGCTCATCCAGCGCCCGACGACAGAGCGGAGCG 2212
DB 4036 TGACAGACTCAAAATATGCATTTGGGAATCTTCAAGCACACCAAGATAGAGTGAATCAG 4095
QY 2213 AGCTGGTGAACCAAGATCATCGAGAGCTGATCAAGAAGAGAGAGTGTACCTGAGCTGG 2272
DB 4096 AGTTAGTCAGTCAAAATAATAGAGCAGTTAAATAAAAAAGAAAAAGTCTACCTGGCATGG 4155
QY 2273 TGCCCGCCCAAGGCGATCGCGGCGCAAGCAGCAGCATCCACAAGCTGGTGACCAAGGCA 2332
DB 4156 TACCAGCACACAAGGAATTTGGAGAAATGAACAAGTAGATGGTGTGCTGCTGCTGAA 4215
QY 2333 TCCGCAAGGTGCTGTCTCTCGACGGCATCGAT 2364
DB 4216 TCAGGAAAGTACTATTTTATAGATGAATAGAT 4247

RESULT 7

US-08-388-353-1
; Sequence 1, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: McPhee, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; APPLICANT: Cooper, David
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388.353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9709 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 44.7%; Score 1100.4; DB 3; Length 9709;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

QY 14 TGCGGAGGCGCATGAGCCAGCCACCA---GCGCAACATCCTGATGTCAGCGCAGCAACT 70
DB 1877 TGCGTGAAGCAATGAGCCAAAGTACCAATCCAGCTACCATTAATGATACAGAAAAGGCAATT 1936
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGCTTCAACTGCGCAAGGAGGCGGCACATCGCCC 130
DB 1937 TTAGGAACCAAGAAAGACATGTTAAGTGTTCAATTTGTGCAAGAGGCGCATAGCCA 1996
QY 131 GCAACTGCGGCGGCCCCGCAAGAAAGGGCTGCTGGAAGTGCAGGAGGCGGCCACACAGA 190
DB 1997 AAAATTGACGGGCGGCTAGGAAAAGAGGCTGTTGAAATGTGAAAGGAAGGACACCAAA 2056
QY 191 TGAAGGACTCACCGAGCGCCAGCCCAACTTCTCCCGAGGAGGACCTGGCCCTTCCCCCAGG 250
DB 2057 TGAAGAGATTACTGAGAGACAGGCTAA-TTTTATGGAAGATCTGGCCCTTCCCACAG 2115
QY 251 GCAAGGCGCCGAGTTCGCCAGCGAGCAGAAACCGCGCAACAGCCCAAGCCAGCGCGGAGC 310
DB 2116 GGAAGGCGAGGATTTCTTCAGAGCAGACAGAGCCACAGCCCAACAGAGAGAGC 2175
QY 311 TGAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGGCGCGAGCGCCAGGGCA 364
DB 2176 TTCAGGTTTGGGAAGAGACAACAACTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGAA 2235
QY 365 -----CCCTGAATTCGCCAGATCACCTGTGGCAGCGCCCTCTGGTGAGGATCAAGG 418
DB 2236 CTGTATCCTTTAGTTCCTTCAGATCACTCTTGGCAGCGACCCCTCGTCACATAAAGA 2295
QY 419 TGGCGCCAGATCAAGAGGCGCCTGTGACACCGCGCGCGGCGGCGGCGGCTGTGAGG 478
DB 2296 TAGGGGCAATTAAGAGAGCTCTATTAGATACAGAGCAGATGATACAGTATTAGAAG 2355
QY 479 AGATGAGCCTGCGCGGCAAGTGGAGCCCAAGATGATCGCGGGGATCGCGGGCTTCATCA 538
DB 2356 AAATGAAATTTGCCAGGAAGATGGAACCAAAAATGATAGGGGAAATTGGAGGTTTATCA 2415
QY 539 AGGTGCGCCAGTAGGACACATCTGTATCAGATCTGCGCAAGAGGCGCATCGGCACCG 598
DB 2416 AAGTAGACAGTAGATGATCAGATCTATAGAAATCTGCGGACATTAAGCTATAGGTACAG 2475
QY 599 TGCTGATGCGGCCACCCCGTGAACATCATCGCGCCCAACATGCTGACCCAGCGTGCGCT 658
DB 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACATGAGTGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCCATCAGAGCCGCTGCGCGGCTGAAGCTGAAGCCGGCA 718
DB 2536 GCACCTTTAAATTTTCCCAATTAGTCTATTGAGACTGTACAGTAAATTTAAAGCCAGGAA 2595
QY 719 TGGACGCGCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCG 778
DB 2596 TGGATGCGCCCAAGGTGAACAGTGGCCCTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAG 2655
QY 779 CCATCTCGAGAGATGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 838
DB 2656 AAATTTGTACAGAAATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2715
QY 839 ACAACACCCCGGTTGCGCCATCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 898

DB 2716 ACAATACTCCAGTATTGCGCATAAAGAAAAGACAGTACTAAATGAGAGAAAATTAGTAG 2775
QY 899 ACTTCGCGGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCGAGCTGGGATCCCC 958
DB 2776 ATTTAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCA 2835
QY 959 ACCCGCGCGGCTGAAGAAGAAAGACGCTGACCGCTGGAGTGGGGGACGCGCTACT 1018
DB 2836 ATCCTGCAGGTTAAACAGAAAATCAGTAACAGTACTGGATGTGGCGATCATATT 2895
QY 1019 TCAGCGTGGCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCACCATCCCCAGCATCA 1078
DB 2896 TTTCAAGTTCCTTAGATAAAGACTTTCAGGAAGTATACATGCTTACCATACCTAGTATAA 2955
QY 1079 ACAACGAGACCCCGGATCCGCTACCACTACACAGTCTGCCCGCAGGCTGGGAAGGCA 1138
DB 2956 ACAATGAGACACGAGGATTAGATATCAGTACATGCTTCCACAGGATGGAAGGAT 3015
QY 1139 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCTCCGGAGCCCTTCCGGCGCCGCAACC 1198
DB 3016 CACCAGCAATATTCAGTGTAGCATGACAAAATCTTAGAGCCTTTAGAAAACAAAATC 3075
QY 1199 CCGAGATCGTGTACTACCA-----GGCCCCCTGTAGTGGGCGAGCAGCTGGAGATCG 1252
DB 3076 CAGACATAGTCACTATCAATACATGATGATTTGTATGATAGGATCTGACTTAGAAATAG 3135
QY 1253 GCAGACACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGGCTTCACCA 1312
DB 3136 GGAGCATAGAACAAAATAGAGAACTGAGAACATCTGTTAGGTGGGATTTACCA 3195
QY 1313 CCGCCGACAAAGACACAGAAAGGAGCGCCCTTCTTGTGATGGGTAGGCTGACGCTGACCC 1372
DB 3196 CACCAGAAAACATCAGAAAGAACCTCCATTCTTGGATGGGTTATGAACCTCATC 3255
QY 1373 CCGACAAAGTGGAGCTGCAGCCCATCGAGCTGCCGAGAGAGAGAGCTGACCGCTGAACG 1432
DB 3256 CTGATNAATGGACAGTACAGCCTATAGTGTCCAGAAAAGACAGCTGGACTGTCAATG 3315
QY 1433 ACATCCAGAAAGCTGGTGGCAAGCTGAACTGGGCGACGAGATCTACCGGCGATCAAGG 1492
DB 3316 ACATACAAATTAAGTGGGAAAATTGAATTTGGGCAAGTCAAGATTTATGCGAGGATTAAG 3375
QY 1493 TGGCGCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGCCCTGACCGACATCGTGGCCCTGA 1552
DB 3376 TAAGGCAATTTATGTAACTTCTTAGGGCAACCAAAAGCACTAACAGAAAGTAGTACCCTAA 3435
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCCTGCGGAGGCGGCTGCAAG 1612
DB 3436 CAGAAGAGCAGAGCTAGAACTGGCAGAAAACAGGGAGATTTCTAAAAGAACCGGTACATG 3495
QY 1613 GCGTGTACTAGACCCCGCAGCAAGGACCTGGTGGCGGAGATCCAGAAGCAGGCGCCAGCC 1672
DB 3496 GAGTGTATTATGACCCCATCAAAAGACTTATAGCAGAAATACAGAACGAGGGGCAAGGCC 3555
QY 1673 AGTGGACCTACAGATCTACAGGAGCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA 1732
DB 3556 AATGGACATATCAAAATTTATCAAGAGCCATTTAAATAATCTGAAAACAGGAAAATATGCA 3615
QY 1733 AGATGCGACCGGCCACACCAAGACGCTGAAGCAGCTGACCGGAGCGGCTGAGAGATCG 1792
DB 3616 GAATGAAGGGTGGCCACACCTAATGATGTGAACAAATTAACAGAGGAGGAGTACAAAAATAG 3675
QY 1793 CCATGGAGACATCGTGTGCGGCAAGACCCCAAGTTCCCGCTCCCATCCAGAGG 1852
DB 3676 CCACAAAGACATAGTAATATGGGAAAGACTTCTTAATTTAAATTTACCATACAAAGG 3735
QY 1853 AGACCTGGGAGACCTGTTGGACCGCACTTACTGGCAGGCGACCTGATCCCGAGTGGAGT 1912
DB 3736 AAACATGGGAAGCATGTTGGACAGAGTATTGGCAAGCCACCTGATTTCTTGGTGGGAGT 3795
QY 1913 TCGTGAACACCCCGCTCGTGAAGCTGTGGTACAGCTGGGAGAGGAGGAGGAGGAGGAGGAG 1972

Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTTACCAAGTTAGAGAAAGAACCCATAATAG 3855
QY 1973 GCGCCGAGACCTTCTAGTGGACGGCCCGCCGACACCAAGCCG 2032
Db 3856 GAGCAGAACTTTCTATGTAGTGGCGCCCAATAGGGAACCTAAATAGGAAAGCAG 3915
QY 2033 GCTACGTGACCGACCGGGCGGAGAGATCGTGAAGCTGACCGAGACCAACCAACAGA 2092
Db 3916 GATATGTAATGACAGAGAGAGACAAAGTTGTCCCCCTAACGGACACACAAATCAGA 3975
QY 2093 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGTGAACATCG 2152
Db 3976 AGACTGAGTTACAAGCAATTCATAGCTTTGCGAGGATTCGGGATTAGAAGTAACATAG 4035
QY 2153 TGACCGACAGCAGTACGCTTGGGCATCATCCAGGCGCCGACCAAGAGCGAGAGCG 2212
Db 4036 TGACAGACTCACAAATATGCAATTTGGGAATCATTCAGCACACCAACAGATAGAGTGAATCAG 4095
QY 2213 AGCTGGTGAACCAAGATCATCGAGCAGCTGATCAAGAAAGGAGAGTGTACTGAGCTGGG 2272
Db 4096 AGTTAGTCAGTCAATATAGAGCAGTTAATAAAAGGAAAGTCTACCTGGCATGGG 4155
QY 2273 TGCCCGCCACAGAGGATCGGGGCAACGAGCAGATCGACAGCTGGTGAAGAGGCA 2332
Db 4156 TACCAGCACACAAGGAATTTGGAGGAATGAACAAGTAGATGGTTCAGTCTGGAA 4215
QY 2333 TCCGCAAGTCTCTCTCGACGCGCATCGAT 2364
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 8

US-08-488-551B-1

; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P#3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: P#4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: P#0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: P#3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGILIO
; REFERENCE/DOCKET NUMBER: 9606Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-488-551B-1

Query Match 44.7%; Score 1100.4; DB 3; Length 9709;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;
QY 14 TGGCCGAGGCATCAGCAGGCCACCA---CGCGCAACATCCTGTATGTCGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCAGTAACATCCAGTACCATAATGATACAGAAAGCAATT 1936
QY 71 TCAAGGCCCCAAGCGCATCATCAAGTCTTCAACTGGCGGAAGGAGGGCCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGACTCTTAAGTGTTCATTTGGCAAAAGAGGGCACAATAGCCA 1996
QY 131 GCAACTGCCCGCCCCCGGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 190
Db 1997 AAAATTGACGGGCCCTTAGGAAAGAGGCTGTGGAAATGTGAAAGAGAGGACACCAA 2056
QY 191 TGAAGGACTGCACGAGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCCAGG 250
Db 2057 TGAAGATTGTACTGAGAGACAGCGCTAA-TTTTATTAGGAAGATCTGGCCCTTCCACAAG 2115
QY 251 GCAAGGCCCGGAGTTTCCCGAGCAGCAGAGAACCGCGCCCAACAGCCCAACAGCGCGAGC 310
Db 2116 GGAAGGCCAGGGAATTTTCTTCAGAGCAGACAGACCCACAGCCCAACAGAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAACCCCCGAGGAGCGCGCGCCGCGCCAGCGGCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAAAGAGAGGAGCCATAGACAAGAA 2235
QY 365 -----CCCTCAACTTCCCCCAGATCACCTGTGGCAGCGCCCTGTGTGAGCTCAAGG 418
Db 2236 CTGTATCCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGCCCTCTGCACATAAGA 2295
QY 419 TGGCGGCGCAGATCAAGAGGCGCCCTGTGGACACCGCGCGCGAGCACACCTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAAGGAAGTCTATTAGATACAGGAGCAGATGATACAGATTAGAAG 2355
QY 479 AGATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGCGATCGGGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGAATTTGAGGTTTATCA 2415
QY 539 AGGTGCGCAGTACGACCAAGATCCTGTATCGAGATCTCGCGCAAGAGGCGCTCGCACCG 598
Db 2416 AAGTAGGACAGTATGATCAGATCTCATAGAAATCTCGGACATTAAGCTATAGTAGAG 2475
QY 599 TGCTGATCGGCGCCACCGCCCTGTAACATCATCGGCGCGCAACATGCTGACCGAGTGGGT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCGCTGAGCTGAGCGCCGCA 718
Db 2536 GCACCTTAAATTTCCCATTTAGTCTTATGAGACTGTACCAAGATAAATAAGAGCCAGAA 2595
QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTTGACCG 778
Db 2596 TGGATGGCCCAAGTAAACAAATGGCCATTGACAGAGAAATAAAGCAATTAGTAG 2655
QY 779 CCATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAAGATCGCCCGCCGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGGAAGGAAATTTTCAAAATTTGGGCTGAAATTCAT 2715
QY 839 ACACACCCCCGTGTTCGCCATCAAGAGAGAGGACAGCACCAAGTGGCGCAGCTGGTGG 898
Db 2716 ACAATCTCAGTATTTGCCATAAAGAAAAAGACAGTAACTAAATGAGAAAAATTAGTAG 2775

QY	899	ACTTCCCGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCCC	958
DB	2776	ATTTCAGAGAAGCTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTTAGGAATACAC	2835
QY	959	ACCCCGCGGCTTGAAGAAGAAGAGCGTGACCGTGCTGGACGTGGCGACCCCTACT	1018
DB	2836	ATCTCGCAGGGTTAAACAGAGAAAATCAGTAAACAGTACTGGATGTGGCGATGCATATT	2895
QY	1019	TCAGCGTCCCTCGAGCGAGGACTTCGCAAGTACACCGCCTTCACCATCCCCAGCATCA	1078
DB	2896	TTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATATGCGATTACCATACTAGTATAA	2955
QY	1079	ACAGAGAGACCCCGGATCCGCTACCACTACAGTGTGCCCCAGGGCTGGGAAGGCA	1138
DB	2956	ACAATGAGACACCCAGGGATTAGATATCAGTACAATGTGCTTCCACAGGATGGAAAGAT	3015
QY	1139	GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGCGCCCGCAACC	1198
DB	3016	CACCAGCAATATTCAGTGTAGCATGACAAAATCTTAGAGCCCTTTAGAAAAACAAAATC	3075
QY	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTCTAGCTGGCGAGCGACCTGGAGATCG	1252
DB	3076	CAGACATAGTCATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAG	3135
QY	1253	GCCAGACCCGGCCAGATCGAGGAGCTGCGCAAGCAOCTGCTGCGTGGGGCTTCACCA	1312
DB	3136	GGCAGATAGAAAACAAAATAGAGAACTGAGACAACATCTGTTGAGGTGGGATTTACCA	3195
QY	1313	CCCCGACAGAGACACCAAGGAGCCCCCTTCTGTGTGATGGCTACGAGCTGCACC	1372
DB	3196	CACCAGCAAAAACATCAGAAAGAACCTTCCATCTCTTTGGATGGGTATGAACTCCATC	3255
QY	1373	CCGACAAGTGGACCGTCGACGCCCATCGAGCTGCCCGAAGAGAGCTTGACCGCTGAACG	1432
DB	3256	CTGATAATGGACAGTACAGCCTATAGTGTGCCAGAAAAGACAGCTGGACTGCTCAATG	3315
QY	1433	ACATCCAGAAGCTGTGTGGCAAGCTGAACTGGCGCAGCCAGATCTACCCCGGCATCAAGG	1492
DB	3316	ACATACAGAAATTTAGTGGGAAAATTTGAATGGCGAAGTCAGATTTATCGCAGGGATTAAAG	3375
QY	1493	TGGCCAGCTGTCAAGCTGCTCGCGCGGCCAAGGCCCTCACCAGCATCTGTGCCCTTGA	1552
DB	3376	TAGGGCAATTTATGTAACCTTCTAGGGGAACCAAGCACTAACAGAAAGTATACCACTAA	3435
QY	1553	CCGAGGAGCGCGAGCTGGAGCTGGCGAGAACCGGGAGATCCTCGCGAGCGCGTGCACG	1612
DB	3436	CAGAAGAAGCAGACTAGACTGGCAGAAAACAGGGAGATTCTTAAAGAACCCGTACATG	3495
QY	1613	CGGTGTACTACGACCCCAAGAGGACCTGGTGGCCGAGATCCAGAAGCAGGGCCACGACC	1672
DB	3496	GAGTGTATTTATGACCCTCAAAAGACTTAATAGCAGAAATACAGNAGCGGGCAAGGCC	3555
QY	1673	AGTGGACCTACCAGATCTACCAGGAGCCCTTCAGAACTGAAACACCGGCAAGTACGCCA	1732
DB	3556	AATGGACATATCAAAATTTATCAGAGCCATTTAAAATCTTGAAAACAGGAAAATATGCAA	3615
QY	1733	AGATGGCCACCGCCACACCAACGACGCTGAAGCAGCTGACCGAGGCGCTGCAGAAGATCG	1792
DB	3616	GAATGAAGGTGCCACACTATGATGTGAACATTTAACAGAGCGAGTACAAAAATAG	3675
QY	1793	CCATGGAGAGCATGTGTGATCTGGGCAAGACCCCAAGTTCGCCCTGCCCATCCAGAAAG	1852
DB	3676	CCAGAAAAGCATAGTAATATGGGAAAAGACTCTCTAAATTTAAATTTACCCATACAAAAG	3735
QY	1853	AGACCTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGATCCCGAGTGGGAGT	1912
DB	3736	AAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGATTTCTTGAAGTGGAGT	3795
QY	1913	TCGTGAACACCCCCCTTGGTGAAGCTGTGTACCAAGCTGGAGAAGGCCCATCATCG	1972
DB	3796	TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAGAAAGAACCCATATAG	3855
QY	1973	GCGCGAGACCTTCTACGTGGAGCGCGCCGCAACCCGCGAGACCAAGATCGGCAAGCGCG	2032

```

RESULT 9
US-09-309-572-15
; Sequence 15, Application US/09309572
; Patent No. 6440730
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral hybrid vectors pseudotyped with LCMV
; FILE REFERENCE: P50489
; CURRENT APPLICATION NUMBER: US/09/309,572
; CURRENT FILING DATE: 1999-05-11
; EARLIER APPLICATION NUMBER: DE 198 56 463
; EARLIER FILING DATE: 1998-11-26
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 9709
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus type 1
US-09-309-572-15

```

Query Match 44.7%; Score 1100.4; DB 4; Length 9709;
Best Local Similarity 68.0%; Pred. No. 1.7e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

Db	2027	AAATGAAATTTGCCAGGAGATGGAAACCAAAAATATGATAGGGGGAATTTGGAGGTTTATFCA	208
Qy	539	AGGTGCGCCAGTACGACACAGATCCTGATCGAGATCTGCGCAAGAAGGCCATCGGCACCG	598
Db	2087	AAGTAGACASTATGATCAGATACTCATASAAATCTGCGGCATTAAGCTATAGGTACAG	2146
Qy	599	TGCTGATGCGCCCCACCCCGTGAACATCATCGGCCGCAACATGCTGACCAGCTGGGCT	658
Db	2147	TATTAGTAGGACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGCT	2206
Qy	659	GCACCCTGAACCTTCCGCATCAGCCCCATCGAGACCGTGCCCGTGAAGCTGAAGCCCGCA	718
Db	2207	GCATTTAAATTTTCCCATTTAGTCTTATTGAGACTGTACCAAGTAAATTTAAAGCCAGGAA	2266
Qy	719	TGGAGCGCCCAAGGTGAAGCTAGCGTGGCCCTGCACCGAGGAGAAGATCAAGGCCCTGACCG	778
Db	2267	TGGATGCCCCAAAGTTTAAACAATGGCCATTGCAAGAGAAGAAATAAAGACATTTAGTAG	2326
Qy	779	CCATCTCGAGAGAGATGGAGAAGGGGCAAGATCACCAAGATCGGCCCGCAGAACCCCT	838
Db	2327	AAATTTGTACAGAAATGGAAAAGGAGGAAAAATTTCAAAATTTGGCCTCGAAAATCCAT	2386
Qy	839	ACAACACCCCGTTTGGCCATCAGAAGAAGACAGACACAAAGTGGCGCAAGCTGGTGG	898
Db	2387	ACAATAGTCCAGTATTTGGCATTAAGAAAGAAAAGACAGTACTAAATGGAGAAAAATTAGTAG	2446
Qy	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTCTGGGAGGTGCAGCTGGGCACTCCCC	958
Db	2447	ATTTCAGAAACTTATAGAGACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCAAC	2506
Qy	959	ACCCCGCGCGCTTGAAAGAAAGAGCGTGACCGTGTGCGACTGGCGCAGCCCTACT	1018
Db	2507	ATCCTGCGAGGTTTAAACAGAAAAATCAGTAACAGTACTGGATGTGGCGCATGCATATT	2566
Qy	1019	TCAGGTGCCCCGTGGAGGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA	1078
Db	2567	TTTCAGTTCCTTAGATAAAGACTTCAGGAAGTACTACTGCATTTACCATACCTAGTATAA	2626
Qy	1079	ACAACGAGACCCCGCATCCGCTACCACTACAGTCAACAGTGCTGCCCGAGGCTGGAGGCA	1138
Db	2627	ACAATGAGACACCGAGGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGAT	2686
Qy	1139	GCCCGACATCTTCCAGACGACATGACCAAGATCCTGGAGCCTTCCGCGCCGCAACC	1198
Db	2687	CACCAGCAATTTCCAGTGTAGCATGACAAAAATCTTAGAGCCTTTAGAAAACAAAATC	2746
Qy	1199	CCGAGATGCTGATCTACCA-----GGCCCCCTGTAGCTGGGCGACGACCTGGAGATCG	1252
Db	2747	CAGACATAGTCACTCATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAG	2806
Qy	1253	GCNAGCACCGCCAGATCTGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTTCACCA	1312
Db	2807	GGCAGCATAGAACAAAAATAGAGGAACCTCAGACACATCTGTTGAGTGGGGATTTACCA	2866
Qy	1313	CCCCGACAAAGACACCAAGAGGACCCCTTCTCTGTGATGGGCTACGAGCTGCACC	1372
Db	2867	CACCAGCAAAAACATCAGAAAGAACCTCCATCTCTTTGATGGGTTATGAATCCCATC	2926
Qy	1373	CCGACAAGTGGCCGTGACGCCCATCGAGCTGCCCGGAGAAGAGAGCTGGACCGTGAACG	1432
Db	2927	CTGATAATGGCAGTACAGCCTATAGTCTGCCAGAAAAAGACAGTGGACTGCTCAATG	2986
Qy	1433	ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGCATCTACCCCGGCATCAAGG	1492
Db	2987	ACATACAGAAATTTAGTGGGAAAAATTTGAATTTGGCGAAGTCAGATTTATGACGGGATTAAG	3046
Qy	1493	TGGGCCAGCTGTCCAAGCTGCTGCGCGGCGCCAGGGCCTGACCGACATCTGTGCCCTGGA	1552
Db	3047	TAAGGCATTTATGTTAACTTCTTAGGGGAACCAAGACACTACAGAGTAGTACCATCAA	3106
Qy	1553	CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCCTGCGGAGCGCGTGCACG	1612
Db	3107	CAGAAGACGAGCTTAGAAGCTGGCAGAAAAACAGGAGAGATTTCAAAGCAACCGTACATG	3166

Qy	1019	1CACGCTGCCCTTGGACGAGGACATCCGAGAGAACACGCTCTACCATCCCCAGCATCA	1071
Db	2567	TTTTCAGTCCCTTAGATAAAGACTTCAGAAAGTACTGCAATTTACCATACCTAGTATAA	2626
Qy	1079	ACAAAGAGACCCCGGCATCCGCTACCAAGTGTCTGCCCCAGGGCTGGAAGGCCA	1138
Db	2627	ACATGAGACACACAGGATTAGATATCACTAGTACAATGTCTCCACAGGATGGAAAGAT	2686
Qy	1139	GCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGCGCCCGCAACC	1198
Db	2687	CACGACCAATATTCCAGCTGTAGCATGACAAAAAATCTTAGAGCCTTTTAGAAAAAATA	2746
Qy	1199	CGAGATCGTGATCTACCA-----GGCCCCCTCTAGCTGGCGAGCAGACCTCGAGATCG	1252
Db	2747	CAGCATAGTCTCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAG	2806
Qy	1253	GCCAGCACCGGCCAAGACTCGAGGAGCTGCGCAAGCACCTGTGCGCTGGGGCTTTCACCA	1312
Db	2807	GGCAGCATAGAACAAAAATAGAGGAACCTGAGACACATCTGTTGAGGTGGGGATTTACCA	2866
Qy	1313	CCCCCGACAGACAGCAGAGAGGAGGCCCCCCCTTCTCTGTGATGGGCTACAGCTGCACC	1372
Db	2867	CACGACAAAAACATCAGAAAGAACCTTCCATTCTTTGGATGGGTTATGAACCTCCATC	2926
Qy	1373	CCGACAAGTGGACCGTCTGAGCCCATCGAGCTGCCCGAAGAGAGACTGGACCGCTGAACG	1432
Db	2927	CTGATAANTGGACAGTACAGCCTATAGTCTGCCAGAAAAGACAGACTGGACTGTCAATG	2986
Qy	1433	ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCAGCCAGATCTACCCCGGCATCAAGG	1492
Db	2987	ACATACAGAAATTTAGTGGGAAAAATTTGAATTTGGCGAAGTCAGATTTATGACGGGATTA	3046
Qy	1493	TGGCGCAGCTGTCAGAGTCTCGCGGGCGCCAAAGGCCCTGACCGACATCTGTGCCCTCGA	1552
Db	3047	TAGGCCAATTTATGTAACTTCTTAGGGGAACCAAGAGACTACAGAGATGTATCCACTTAA	3106
Qy	1553	CCGAGGAGCCGAGCTGGAGCTGGCCGGAAGACCGCGAGATCCTGCGGAGCGCGTGCACG	1612
Db	3107	CAGAAGACGAGCTTAGAAGCTGGCAGAAAACAGGAGAGATTTCAAGAAACACCGTACATG	3166

```
QY 1613 CGGTGTTACTACGACCCAGCAAGGACCTGTGGCGGAGATCCAGAAGCAGGCGCCACGACC 1672
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 3167 GAGTGTATTATGACCCATCAAAAGACITTAATAGCAGAAATACAGAAGCAGGCGCAAGGCC 3226
QY 1673 AGTGGACCTACCATCTACGAGGAGCCCTTCAGAAGCTGGAAGCAGGCGCAAGTACGGCA 1732
    ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 3227 AATGGACATATCAATTTATCAAGGCAATTTAAAAATCTGAAACACAGGAAATATGCAA 3286
QY 1733 AGATGGCGACCGCCACCAACCAACGAGTGAAGCAGCTGACCGAGGCGCTGCAGAGATCG 1792
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3287 GAATGAAGGTGCCACACACTAATGATGGAACAAATTAACAGAGCGATACAAAAATAG 3346
QY 1793 CCATGGAGAGACCTGTGATCTGGGCGAGACCCCAAGTTCCTGGCTGCCCTCCAGAAAG 1852
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3347 CCACAGAAAGCATAGTAAATATGGGAAAGACTCCTAATTTAAATACCCATACAAAGG 3406
QY 1853 AGACCTGGGAGACCTGTGGACGACACTACTGGCAGGCGCAAGTACCTGATCCCGAGTGGAGT 1912
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3407 AAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGGGAGT 3466
QY 1913 TCGTGAACACCCCCCTCGTGAAGCTGTGTACCAAGCTGGAAGAGAGCCCATATCG 1972
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3467 TTGCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTTGAAGAAAGAACCCATAATAG 3526
QY 1973 CGCGCGAGACCTTCTAGCTGGACGCGCGCCCAACCCGAGACCAAGATCGGCAAGGCGC 2032
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3527 GAGCAGAAACTTTCTATGTAGATGGGCGAGCCAAATAGGAAACTTAATTAGAAGAGCAG 3586
QY 2033 GCTAGTGGACGACCGGCGCGGAGAGATCGGTGAGCTGACCGAGACCAACCAACAG 2092
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3587 GATATGAACCTGAGAGAGGAAAGAAAAGTTGTCCCTTAACGAGACACAACAAATCAGA 3646
QY 2093 AGACGAGCTCGAGCCATCCAGCTGGCCCTGAGGAGCAGCGGCGAGCTGACATCG 2152
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3647 AGACTGAGTTACAAGCAATTCATCTAGCTTTCAGGATTCGGGATTAAGAAATAACATAG 3706
QY 2153 TGACCGACAGCCAGTACGCCCTGGGCAATCCAGGCGCCAGCCGACAAAGAGCGAGAGCG 2212
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3707 TGACAGACTCAATATGCAATTTGGAAATCAATCAACACACACAGATGAAGTGAATCAG 3766
QY 2213 AGCTGGTGAACCATCATCCAGCAGCTGATCAAGAGGGAAGGTGTACTGAGCTGGG 2272
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3767 AGTTAGTCAGTCAATATAGACAGTAAATATAAGGAAAGGAAAGTCTACCTGGCATGGG 3826
QY 2273 TGCCCGCCACAGGCGATCGGCGGCAACGAGCAGATCGACAACTGGTGAAGAGGCA 2332
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3827 TACCAGCACAAAGGAATGGAGGAATGAACAGTAGATGGGTGTCAGTCTGTGGA 3886
QY 2333 TCCGCAAGGTGCTTCTCGGACGCGCATCGAT 2364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3887 TCAGGAAAGTACTATTTTAGATGGAATAGAT 3918
```

RESULT 11

US-08-848-760B-33

; Sequence 33, Application US/08848760B

; Patent No. 6248721

; GENERAL INFORMATION:

; APPLICANT: Chang, Lung-Ji

; TITLE OF INVENTION: Animal Model For Evaluation Of Vaccines

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

; STREET: 2421 N.W. 41st Street, Suite A-1

; City: Gainesville

; STATE: Florida

; COUNTRY: United States of America

; ZIP: 32606

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/848,760B

; FILING DATE: 25-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/838,702

; FILING DATE: 09-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: PACE, DORAN R.

; REGISTRATION NUMBER: 38,261

; REFERENCE/DOCKET NUMBER: CNG-100C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (352) 375-8100

; TELEFAX: (352) 372-5800

; INFORMATION FOR SEQ ID NO: 33:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 12494 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: other nucleic acid

; DESCRIPTION: /desc = "DNA"

; SEQUENCE DESCRIPTION: SEQ ID NO: 33:

US-08-848-760B-33

```
Query Match      44.7%; Score 1100.4; DB 4; Length 12494;
Best Local Similarity 68.0%; Pred. No. 1.8e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;
```

QY 14 TGGCGGAGGCGCATGAGCCAGGCGCACCA---GGCCCAACATCTCTGATGCGCGCAGCAACT 70

Db 1548 TGGCTGAGAGCAATGAGCCCAAGTAACAATCCAGCTACCATATGATACAGAAAGCAATT 1607

QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGGCCACATCGCCC 130

Db 1608 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTGGCAAGAGGACATAGCCA 1667

QY 131 CCAACTGCGCGCCCCCGCAGAGAGGCTGCTGGAAGTGGCGCAAGAGGCGCCACCAGA 190

Db 1668 AAAATTGCGAGGCCCCCTAGGAAAAGGCTGTGGAAATGTGGAAGGAAGGACACCAAA 1727

QY 191 TGAAGGACTGCGAGCGAGCCAGCCCAACTTCTTCCGCGAGGACCTTGCTTCCCCAGG 250

Db 1728 TGAAGATTGTACTCAGAGACAGGCTAA-TTTTATTAGGAAGATCTGGCTTCCCAACAG 1786

QY 251 GCAAGGCCCGGAGTTCCTCCAGCGAGCAGAACCGCCCAACAGCCCCACCGCCGAGC 310

Db 1787 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACCAAGCCCAAGCCCCACCAAGAGAGC 1846

QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGGCGCCGAGCGCCAGGGCA 364

Db 1847 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAAAGCAGAGGCGCATAGACAAGGAA 1906

QY 365 -----CCCTGAACCTTCCCCAGATCACCTGTGGCAGCGCCCTGTGTGACATCAAGG 418

Db 1907 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGGAGCCCTCTGTCACAATAAGA 1966

QY 419 TGGCGGCGCAGATCAAGGAGGCCCTGTGGACACCGCGCGCCAGACACCGCTGTGGAGG 478

Db 1967 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 2026

QY 479 AGATGAGCCTGCCCGGAAGTGGAAAGCCCAAGATGATCGGGGGCATCGCGGCTTCATCA 538

Db 2027 AAATGAATTTGCCAGGAAGATGGAAACCAAAATGATAGGGGGAATTTGAGGTTTTATCA 2086

QY 539 AGGTGCGCGCAGTACGACCAAGATCCTGTGATCGAGATCTCGGCAAGGAAGCCATCGCACCG 598

Db 2087 AAGTAGGACATATGATCAGATACATAGAAATCTCGGACATTAAGCTATAGTAGAG 2146

QY 599 TGCTGATCGGCCCCACCCCGCTGAACATCATCGGCGCGCAACATGTCGCCAGCTGGCT 658

Db 2147 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGCT 2206

ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 1581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: -
LOCATION: 1..15581
OTHER INFORMATION: /note= "pNUnSg11"
US-08-646-538-35

Query Match 44.7%; Score 1100.4; DB 3; Length 15581;
Best Local Similarity 68.0%; Pred. No. 1.8e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;
Qy 14 TGGCGAGGCGCATGAGCCAGCCACCA---GCGCCAAACATCCTGATGCGAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCCAAGTAACAAATCCAGCTACCCTATGATGATACAGAAAGGCAATT 1936
Qy 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGGAGGGCCACATCGGCC 130
Db 1937 TTAGAACCAAGAAAGACTGTTAAGTGTTCATATTTGGCAAGAAAGGGCACCATAGCCA 1996
Qy 131 GCAACTTCGCGGCCCCCGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
Db 1997 AAAATTGACGGCCCTTAGGAAAAAGGCTCTTGGAATGTGGAAGAAAGGACACACAAA 2056
Qy 191 TGAAGGACTGCACGAGCGCCAGGCGCAACTTCTTCCGGGAGAGCTGGCCCTTCCCGCAGG 250
Db 2057 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTGTAGGGAAGATCTGGCCCTCCACAAAG 2115
Qy 251 GCAAGGCGCGGAGTTCCCGACGAGCAGAACCGCGCCACAGCCACCGCCAGCGCGAGC 310
Db 2116 GGAAGCCAGGGAATTTCTTCAGAGCAGACAGCCACAGCCACAGCCCGCCACAGAGAGAGC 2175
Qy 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGGCGCCGAGCGCCAGGSCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAAACAATCCCTCTCAGAGAGAGAGCCGATAGACAAAGAA 2235
Qy 365 -----CCCTGAACCTTCCCGCAGATCACCTCTGGCAGCGCCCTGGTGAGCATCAAGG 418
Db 2236 CTGTATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACAATAAGA 2295
Qy 419 TGGCGGCGCAGATCAAGAGGCGCTGCTGGACACCGGCGCGCAGACACCGTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGATTAGAAAG 2355
Qy 479 AGATGAGCCTGCCCGGAAGTGAAGCCCAAGATGATCGCGGCGATCGGCGGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATGATAGGGGGAATTGGAGGTTTTATCA 2415
Qy 539 AGGTGCGCCAGTACGACCATCTGATCGAGATCTCGGGCAAGAGCCATCGCGACCG 598
Db 2416 AAGTAGACAGATGATCAGATCTATAGAAATCTCGGACATCTCGGACATTAAGCTATAGGTACAG 2475
Qy 599 TGCATGATCGGCGCCCGCCGCTGAACATCATCGCGCCGCAACATCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGACCTACACCTGTCAACATAATTGGAAGAAATCTCTTACTCAGATTTGGCT 2535
Qy 659 GCACCTTGAATTTCCCATCAGCCCCATCGAGACCGTCCCGTGAAGCTGAAGCCCGGCA 718
Db 2536 GCACCTTTAAATTTTCCCATATTGCTTATGAGACTGTACCAAGTAAATTAAGCCAGGAA 2595
Qy 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGAGAGATCAAGGCCCTGACCG 778

Db 2596 TGGATGGCCCAAAAGATTAAACAATGCCATTGCACAGAGAAAAAATAAAGCATTAGTAG 2655
Qy 779 CCATCTGCGAGGAGATGGAGAAGGCGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 2656 AAATTTGTACAGAAATGGAAGAGGAAAAAATTTCAAAATTTGGCGCTGAAAAATCCAT 2715
Qy 839 ACAACACCCCGTGTTCGCCATCAAGAGAAGACAGACACCAAGTGGCCCAAGCTGGTGG 898
Db 2716 ACAATACTCCAGTATTGGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 2775
Qy 899 ACTTCCGCGAGCTGAACAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958
Db 2776 ATTTCCAGAGAATTAATTAAGAACTCAAGATTTCTGGGAAGTTCAATTAGGAATACCA 2835
Qy 959 ACCCGCGCGCTGAAGAGAAGAGAGGCTGACCGTGTGAGCGTGGCGGACGCGCTACT 1018
Db 2836 ATCTCTGCAGGGTTAAATCAGAAAAAATCAGTAACAGTACTGGATGTGGCGATGCATATT 2895
Qy 1019 TCAGCGTGCCTGACGAGGACTTCCGCAAGTACACCGCTTCACCATCCCCCAGCATCA 1078
Db 2896 TTTCAAGTTCCCTTAGATAAAGACTTTCAGGAAGTATACTGCATTTACCATACCTAGTATA 2955
Qy 1079 ACAACGAGACCCCGCGCATCCGCTACCAAGTACAAGTGTGCTCCCGAGGCTGGAAGGCA 1138
Db 2956 ACAATGAGACACCGAGGATTAGATATCAGTACAATGTGCTTCCAGAGGATGGAAGGAT 3015
Qy 1139 GCCCGCAGCATTTCCAGAGCAGCATGACCAAGATCCTTGAGCGCTTTCGCGCGCCCAAGC 1198
Db 3016 CACCAAGCAATATTCAGTGTAGCATGACAAAAATCTTAGAGCCTTTAGAAAAACAATC 3075
Qy 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTCTAGTGGCGAGCGACCTGGAGATCG 1252
Db 3076 CAGACATAGTCACTCATCAATACATGGATGATTTGTATGAGTACTGACTTAGAAATAG 3135
Qy 1253 GCCAGCACCGGCGCAAGATCGAGGAGCTGGCAAGCACCTGCTGGCTGGGCTTCACCA 1312
Db 3136 GGCAGCATAGAACAAAAATAGAGGAATCGAGACAACATCTGTTAGGTGGGATTTACCA 3195
Qy 1313 CCCCCGCAAGAGACACCAAGAGGAGCGCCCTTCTGTGGATGGCTACGAGCTGCACG 1372
Db 3196 CACCAAGAAAAAATCAGAAAGAACCTCCATTTCTTTGGATGGTTATGAACCTCAATC 3255
Qy 1373 CCGAAGCTGGACCTGCGAGCCATCGAGCTGCCGAGAGAGAGCTGAGCCCGTGAAGC 1432
Db 3256 CTGATAATGGACAGTACAGCTTATGCTGCTCCAGAAAGAGCAGCTGGACTGTCAATG 3315
Qy 1433 ACATCCAGAAAGCTGGTGGGCAAGCTGAACTGGGCCAGCAGATCTACCCCGGCATCAAG 1492
Db 3316 ACATACAGAAATTAGTGGGAAAAATGAAATGGGCAAGTCAGATTTATGACGGGATTAAG 3375
Qy 1493 TGGCCGACCTGTGCAAGCTGCTGCGCGGCGCAAGGCGCTGACCCAGATCTGTCGCCCTGA 1552
Db 3376 TAAGCAATATGTAACCTTCTTAGGGGAACCAAGACTAACAGAAAGTAGTACCATAA 3435
Qy 1553 CCGAGGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGAGCGCCGTCAGG 1612
Db 3436 CAGAAAGCAGAGCTAGAACTGGCAGAAACAGAGGAGATTTCTAAAGAACCGGTACATG 3495
Qy 1613 GCGTGTACTACGACCCCGAGCAAGGACCTGGTGGCGGAGATCCAGAAAGCAGGCGCCAGC 1672
Db 3496 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGCGCAAGGCC 3555
Qy 1673 AGTGACCTACAGATCTTACCAGGAGCCCTTCAAGAACTCTGAAGACCGCGCAAGTACGCA 1732
Db 3556 AATGACATATCAAAATTTATCAAGGCAATTTAAAAATCTGAAAAACAGGAAAAATGCAA 3615
Qy 1733 AGATGCGCACCGCCACACCAAGCAGCTGAAGAGAGCTGACCGAGCGCTGAGAGATCG 1792
Db 3616 GAATGAAGGGTGGCCACACTAATGATGTAACAACTTAACAGAGCGCATACAAAAATAG 3675
Qy 1793 CCATGGAGAGCATCGTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 1852

Db 3676 CCACGAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAAATTAACCCATACAAAAGG 3735
QY 1853 AGACCTGGGAGACCTGGTGGACGACTACTGGCAGGCCACCTGGATCCCGGAGTGGAGT 1912
Db 3736 AAACATGGGAAGCATGGTGGACAGATATATGGGAAGCCACCTGGATTCCTGAGTGGGAGT 3795
QY 1913 TCCTGAACACCCCCCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCCATCATCG 1972
Db 3796 TTGTCATACCTCCCTTAGTAGAGTATAGTACCAAGCTGGAGAGAACACCCATATAG 3855
QY 1973 GCGCCGAGACCTTCTACGTGGAGCGGCCGCCCAACCGCAGAGACCAAGATGGCAAGCGCG 2032
Db 3856 GAGCAGAAACTTTCTATGTAGATGGGCGAGCCCAATAGGGAAACTAAATTAGGAAAGCAG 3915
QY 2033 GCTACGTGACCGAGCGCGGCCGAGAGATCGTGGACCTGACCGAGACCAACCAACAGA 2092
Db 3916 GATATTAAGTACAGAGGAGAACAAAAGTTGTCCCTTAACGGACACACAAATCAGA 3975
QY 2093 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCG 2152
Db 3976 AGACTGAGTACAGCAATTCATCTAGCTTTCAGGATTCGGGATTAGAGTAACATAG 4035
QY 2153 TGACGACAGCAGTACGCGCTGGGATCATCGAGGCCCGCAGCCGACAGAGCGAGCG 2212
Db 4036 TGACAGACTCACAAATATGCAATGGGAATCATTCAGGACCAACAGATAGAGTGAATCAG 4095
QY 2213 AGCTGGTGAACAGATCATCGACGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2272
Db 4096 AGTTAGTCAGTCAAAATAGAGCAGTAAATAAAAGGAAAGTCTACTGGCATGGG 4155
QY 2273 TGCCCGCCCAAGAGGCGATCCGCGCAACGACGACATCGACAAGCTGGTGAGCAAGGGCA 2332
Db 4156 TACGAGCACAAAGGAAATGGAGGAATGAACAAGTAGATGGTGGTCTGCTGGAA 4215
QY 2333 TCCGCAAGTGTCTTCTCGGAGCGCATCGAT 2364
Db 4216 TCAGGAAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 13

US-09-503-222-35
; Sequence 35, Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Galtanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/503,222
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..1581
; OTHER INFORMATION: /note= "pNLS611"
US-09-503-222-35

Query Match 44.7%; Score 1100.4; DB 4; Length 1581;
Best Local Similarity 68.0%; Pred. No. 1.8e-172;
Matches 1614; Conservative 0; Mismatches 736; Indels 22; Gaps 5;

QY 14 TGCCCGAGGCCATGAGCCAGGCCACCA---CGGCCACATCCTGATCGAGCGCAGCAACT 70
Db 1877 TGCTGAAGCAATGAGCCAAAGTAACAAATCCAGTACCAATTAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGCCCCCAAGGCGATCATCAAGTGTCTCAACTGCGCAAGGAGGCGCAATCGCCC 130
Db 1937 TTAGGAACCAAAAGAGACTGTTAAGTGTTCAAATTTGTGCAAGAGAGGCGACATAGCCA 1996
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGGAGAGGCGCCACAGA 190
Db 1997 AAAATTGCAGGCGCCCTAGGAAAAAGGGCTGTTGGAATGTGGAAGAGGAGGACACCAA 2056
QY 191 TGAAGACTCGACGAGCGCCAGGCCCAACTTCTTCCCGAGGAGCTGCGCTTCCCCCAGG 250
Db 2057 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTTTAGGGAAGATCTGCGCTTCCCAAG 2115
QY 251 GCAAGGCCCGCGAGTTTCCCGAGCGAGCAAGCGCGCAACAGCCGCCACCCAGCGCGCAG 310
Db 2116 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGACAGCCACCCAGGAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCAGGGCA 364
Db 2176 TTCAGGTTTGGGGAAGAGACAAACCTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAA 2235
QY 365 -----CCCTGAACCTTCCCCCAGATCACCTGTGGCAGCGCCCTGTTGAGCATCAAG 418
Db 2236 CTGTATCTCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCTGTCACAATAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGCAGCACCGTGTGGAGG 478
Db 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGATTAGAG 2355
QY 479 AGATGAGCCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGGATCGGGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAATATAGGGGAAATTTGAGGTTTTATCA 2415
QY 539 AGTGCCCGCAGTAGCAGCAGATCCTGATCGAGATCTCGGCAAGAGGCCATCGCACCG 598
Db 2416 AAGTAGGACAGATGATGATCAGATCACTATAGAAAATCTCGGACATAAAGCTATAGGTAC 2475
QY 599 TGCTGATCGGCGCCCGCCCGTGAACATCATCGGCCCGCAACATCGTCAACCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTTGAACCTTCCCATCAGCCCCATCGAGACCGTGCCTGCGGCGTGAAGTGAAGCCCGCA 718
Db 2536 GCACCTTTAAATTTTCCCATTAGTCTCTATTGAGACTGTACCAAGTAAATTAAGGACAGAA 2595
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGCCCTGACCG 778
Db 2596 TGGATGCCCAAAAGTTAAACATGGCCCATTGACAGAGAAAAAATAAAGCATTAGTAG 2655

QY	779	CCATCTCGGAGGAGATGGAGAAAGGAGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT	838
Db	2656	AAATTTGTACAGAAAATGGAAAATGGAGGAGAAAATTTCAAAAATTTGGGCCTGAAAATCCAT	2715
QY	839	ACAACACCCCGTTCGCCCATCAAGAAGAGGAGACACCAAGTGGCGCAAGCTGGTGG	898
Db	2716	ACAACTCCAGTATTTGGCCATAAGAGAAAACAGTACTAAATGGAGAAAATTTAGTAG	2775
QY	899	ACTTCGCGAGCTGAACAAGCGCACCCAGGACCTCTGGGAGGTGCAGCTGGGCATCCCCC	958
Db	2776	ATTTCAGAGAACCTTAATAAGAGACTCAAGATTTCTGGGAAGTTCAATTAGGATATCCAC	2835
QY	959	ACCCGCGCGCTGAAGAAGAAAGAGAGCGTGACCGTGCTGGAGCTGGGGAGCGCTTACT	1018
Db	2836	ATCCTGCAAGGTTTAAACACAGAAAAATCAGTAACAGTACTGGATGTGGCGATGCATATT	2895
QY	1019	TCAGCGTGCCTGGAGGAGGACTTCCGCAAGTACACCGCTTCCACCATCCCGAGCATCA	1078
Db	2896	TTTCAGTTCCTTAGATAAAGACTTCAGGAAGTATATGSCATTTACCATACCTTAGTATAA	2955
QY	1079	ACAACGAGACCCCGGCATCCGCTACCAAGTACAACGCTGCTGCCCGCAGGGCTGGAAAGGCA	1138
Db	2956	ACAATGAGACACCGAGGATTAGATATCAGTACAATGCTCTCCACAGGGATGGAAAGGAT	3015
QY	1139	GCCCCAGCATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCGGCGCGCAAC	1198
Db	3016	CACGACGAATATTCCAGTGTAGCATGACAAAAATCTTAGAGCCTTTAAAAACAAAAATC	3075
QY	1199	CCGAGATGCTGATCTACCA-----GGCCCCCTCTAGCTGGCGAGCGACTGGAGATCG	1252
Db	3076	CAGACATAGTCATCTCAATACATGGATGATTTGTATGATAGGATCTGCATTTAGAAATAG	3135
QY	1253	GCCAGCACCGCGCAAGATCTGAGGAGCTGCGCAAGCACCTGCTCGGTGGGGCTTTCACCA	1312
Db	3136	GGCAGCATAGAACAAAAATAGAGAACTCAGACACATCTGTTGAGGTGGGGATTACCA	3195
QY	1313	CCCCGACAAAGACACAGAGAGAGCCCCCTTCTGTGATGGGTACGAGCTGCACC	1372
Db	3196	CACGACAAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGGTATGAATCCCATC	3255
QY	1373	CCGACAAGTGGACCGTGACGCCATCGAGCTGCCCGAAGAGAGCTGGNACCTGNAACG	1432
Db	3256	CTGATTAATGGACAGTACAGCCCTATAGTGTGCCAGAAAAGGACAGCTGGACGTGCTAATG	3315
QY	1433	ACATCCAGAGCTGTGTGGCAAGCTGAACTGGGGCAGCGCCAGTCTACCCGGGATCAAGG	1492
Db	3316	ACATACAGAAATTAGTGGCAAAATTTGAATTTGGCAAGTCAGATTTATCGCAGGATTAAG	3375
QY	1493	TGGCGCAGCTGTCAAGCTGTCTGCGGGCGCAAGGCCCTGACCGACATCTGTGCCCTTGA	1552
Db	3376	TAAGGCAATATTGTAACCTTCTTAGGGGAACCAAAAGCACTAACAGAAGTAGTACCACTAA	3435
QY	1553	CCGAGAGCGCAGCTGGAGCTGGCCGAGACCGCGAGATCTCTCGCGAGCCCGTGCACG	1612
Db	3436	CAGAGAGCAGAGCTAGAACTTGGCAGAAAACAGGGAGATTCTAAAAAGAACCGGTACATG	3495
QY	1613	GCGTGTACTACACCCACCAAGAGACTTGGTGGCCGAGATCCAGAACGAGGGCCACGACC	1672
Db	3496	GAGTGTATTGACCCATCAAAAGACTTAATAGCAGAAATACAGACGAGGGCGCAAGGCC	3555
QY	1673	AGTGGACCTACAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTAGCCCA	1732
Db	3556	AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAAATATGCAA	3615
QY	1733	AGATGCGCACGCCCCACACACGACGTGAAGCAGCTTGACCGAGCCGTGCAGAAATGCG	1792
Db	3616	GAATGAAGGGTGCCACACACTAATGATGTGAAACAATTAACAGAGGGCAGTACAAAAATAG	3675
QY	1793	CCATGAGAGCATGCTGTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGG	1852
Db	3676	CCACAGAGCATAGTATATGGGGAAAGACTCTCAATTTAAATTTACCCATACAAAAGG	3735
QY	1853	AGACCTGGGAGACCTTGGTGGACCGCACTACTGGCAGGCCACCTGGATCCCCGAGTGGGAGT	1912

Db	3736	AAACATGGGAAGCATGGTGGACAGAGATTATTGGCAAGCCACCTGGATTCTCGTGGTGGAGT	3795
Qy	1913	TCGTGAACACCCCCCCCTGGTGAAGCTGTGTACCACTGGAGAGAGAGCCCATCATCG	1972
Db	3796	TTGTCAATACCCCTCCCTTAGTGAAGTTATGTGTACCACTTAGAGAAAGAACCCATAATAG	3855
Qy	1973	GCGCCGAGACCTTCTACGTGGAGCGGCCGCCCAACCGCGGAGACCAAGATCGGCAAGGCCG	2032
Db	3856	GAGCAGAAACTTTCTATGTAGATGGGCGAGCCAATAGGAAACTTAAATTAGAAAGACGAG	3915
Qy	2033	GCTACGTGACCGACCGGGCGGCGCAGAGAAGATCGTGAAGCTGACCGAGACCAACCAACGAGA	2092
Db	3916	GATATGTAACTGCAGAGGAAGACAAAAAGTTGTCCCTTACCGGACACAAACAATCAGA	3975
Qy	2093	AGACCGAGCTGCGAGGCGCATCCAGCTGGCCCTTCGAGGACAGCGGACGAGGTGAACATCG	2152
Db	3976	AGACTGAGTTACAAAGCAATTCATCTAGCTTTCGAGGATTCGGGATTAGAAGTAAACATAG	4035
Qy	2153	TGACCGACAGCCAGTAGTACGCCCTGGGCGATCATCCAGGCCCGCCGACAGAGCGAGCGG	2212
Db	4036	TGACGAGACTCACAAATATGCAATTTGGGAATCATTTCAAGCACAAACCAGATAAAGTGAATCAG	4095
Qy	2213	AGCTGGTGAACAGAGATCATCGACGAGCTCATCAAGAAGGAGAGAGTGTACCTGAGCTGGG	2272
Db	4096	AGTTAGTCACTCAAAATAATAGACAGTTTAATAAAGAAAGAAAAGTCTACCTGGCATGGG	4155
Qy	2273	TGCCCGCCCCAAGAGGCGATCGCGGCCCAACGAGCAGATCGACAAGCTGGTGAAGCAAGGCCA	2332
Db	4156	TACCAGCACAAAGGAATTTGGAGGAATAATGAACAAGTAGATGGGTGGTCAGTCTCGAA	4215
Qy	2333	TCCGCAAGTGCTGTTCCCTGGAGGCGATCGAT	2364
Db	4216	TCAGGAAAGTACTAATTTTAGATGGAATAGAT	4247

RESULT 14

US-08-944-449-7

Sequence 7, Application US/08944449

Patent No. 5985613

GENERAL INFORMATION:

APPLICANT: KURTH, REINHARD

APPLICANT: BAIER, MICHAEL

APPLICANT: METZNER, KARIN

APPLICANT: WERNER, ALBRECHT

TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing

TITLE OF INVENTION: lymphokine (iSL)" to inhibit the replication of

TITLE OF INVENTION: viruses, particularly of retroviruses

FILE REFERENCE: 8341-7065

CURRENT APPLICATION NUMBER: US/08/944,449

CURRENT FILING DATE: 1997-10-06

EARLIER APPLICATION NUMBER: EP 95113013.2

EARLIER FILING DATE: 1995-08-18

EARLIER APPLICATION NUMBER: DE 195 13 152.5

EARLIER FILING DATE: 1995-04-07

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 7

LENGTH: 9737

TYPE: DNA

ORGANISM: Human immunodeficiency virus type 1

US-08-944-449-7

		Query Match	44.3%	Score 1090.8;	DB 2;	Length 9737;
		Best Local Similarity	67.8%;	Pred. No. 6.5e-171;		
		Matches 1608; Conservative	0;	Mismatches 742;	Indels 22;	Gaps
QY	14	TGCGCCAGGCCATGAGCCAGGCCACCA---	GGGCCAACATCCTCATGCAGCGCGAGCAACT	70		
Db	1884	TGCGTGAAGCCATGAGCCAAAGTAACAAATCCACGTAAACAT	AATGATGCAGAGGCAAAAT	1943		
QY	71	TCAAGGGCCCCAAGCGCATCATCAAGTGCTTTCAACTCGGCAAGSAGGGGCCACATCGCCC		130		

Db 1944 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGCGAAGAAAGGCACATAGCCA 2003
QY 131 GCAACTGCCGCGCCCGCAAGAGGCTGCTGGAAGTGCAGAGAGGCGCACACAGA 190
Db 2004 AATAATGCGAGGCGCCCTAGGAAGAAAGGCTGTGGAGATGTGAGAGGAAGACACACAA 2063
QY 191 TGAAGGACTGACCGAGCGCCAGGCCCAACTTCTTCCGCGAGGACCTGCCCTTCCCCAGG 250
Db 2064 TGAAGATTGCACTGAGAGACAGSCTAA -TTTTTTAGGAAGATCTGSCCTCTCTACAAG 2122
QY 251 GCAGGCGCGGAGTTCCTCAGCGAGCAGACACCGCCACAGCCACAGCGCGAGC 310
Db 2123 GGAAGCGCAGGGAATTTCTTCTAGAGCAGACACAGAGCAACAGCCACACAGAGAGC 2182
QY 311 TGAGGTGCGCGGGG-----ACAACCCCGCAGCGAGGCGCGCGAGCGCGCAGGCA 364
Db 2183 TTCAAGTTTGGGAGGAGAAACACTCCCTCTCAGAGCAGGAGCGCCGATAGACAAGAA 2242
QY 365 -----CCCTGAATTCCTCCAGATCACTCCCTGTGGCAGCGCCCTGGTGAGCATCAAG 418
Db 2243 CTGTATCTCTTAATCTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAATAAGGA 2302
QY 419 TGGCGCGCAGATCAAGGAGCGCTGCTGGACACCGCGCGGAGCAGCAGCGTGTGGAG 478
Db 2303 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGATTAGAAG 2362
QY 479 AGATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATCGGGCGATCGGGCGCTTCATCA 538
Db 2363 AATGAATTTGCCAGGAAATGAAACCAAAATGATAGGGGAAATTTGGAGTTTATCA 2422
QY 539 AGGTGCGCCAGTAGCAGACAGATCTGTAGATCTCGGCGAAGAGCCATCGGCACCG 598
Db 2423 AAGTAAGACAGTAGCATACCTCTAGAAATCTGTGGACATAAGCTATAGGTACAG 2482
QY 599 TGCTGATCGGCGCCACCCCGTGAACATCATGCGCCGCACATGCTGACCCAGCTGGCT 658
Db 2483 TATTAGTAGGACCTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGTT 2542
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGCCTGAAGCTGAAGCCCGGCA 718
Db 2543 GTACTTTAATTTCCCATTTAGTCTTATTGAACTGTACCAGTAAATTAAGCCAGGAA 2602
QY 719 TGAAGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2603 TGGATGCGCCCAAGGTGAAGCAATGGCCATTGACAGAGAAATAAAGCATTAGTAG 2662
QY 779 CCATCTCGAGGAGATGAGAGAGGAGGCAAGATCACCAGATCGGCGCCCGAGNACCCCT 838
Db 2663 AGATATGTACAGAAATGGAAGAGGAGGAAATTTCAAAATTTGGCCCTGAAATCCAT 2722
QY 839 ACAACACCCCGCTTCCCGCATCAAGAGAGGACAGACCAAGTGGCGCAAGCTGGTG 898
Db 2723 ACAATCTCCAGTATTGCTTATAGAGAAAGACACTACTAATGAGAGAACTAGTAG 2782
QY 899 ACTTCGCGAGCTGAACAAGCCACCGAGTCTTGGAGGTGACGCTGGGCATCCCC 958
Db 2783 ATTTCAGAGAACTTAATAAAGAACTCAAGACTTCTGGGAAGTTCAAGTAGGAATACCAC 2842
QY 959 ACCCGCGCGCTGAAGAGAGAGAGCGTGCACCGCTTCCAGCATCCCCAGCATCA 1018
Db 2843 ACCCGCAGGGTTAAAAAGAGGAAATTCAGTAACAGTATTGGATGTGGGTGATGCATCT 2902
QY 1019 TCAGCGTGCCTTGAGCAGGACTTCCGCAAGTACACCGCTTCCAGCATCCCCAGCATCA 1078
Db 2903 TTTTCAGTTCCCTTAGATAAAGACTTTAGAAAGTATATGCTATACCATACCTAGTATAA 2962
QY 1079 ACAACAGAGACCCCGCATCCGCTACAGTACAACTGCTGCCCGAGGCGTGGAGGCA 1138
Db 2963 ACAATGAGACACCGGATTAGATATCAGTACAATGTGCTGCCACAGGGATGGAAGCAT 3022
QY 1139 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCTTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 3023 CACCAGCAATATTCCAAAGTAGCATGACAAAAATCTTTAGAGCCCTTTTAGAAAAACAGAACT 3082

QY 1199 CCAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGAGGACCTGGAGATCG 1252
Db 3083 CAGACATAGTTATCTATCAATACATGATGATTGTATGTAGGATCTGACTTGAATAG 3142
QY 1253 GCCAGCAGCGCGCAAGATCGAGGAGTGGCAAGCAACCTGCTGCGTGGGGCTTCACCA 1312
Db 3143 GGCAGCATAGAACAAAAATAGAGAACTGACAGCATCTCTTGGAGTGGGATTTACCA 3202
QY 1313 CCCCCGACAAGACACCAAGAGAGCCCCCTTCTTCTGGATGGGTACGAGCTGCACC 1372
Db 3203 CACGACAAAAACATCAGAAAGAACCTCCATTTCTTGGATGGGTATGAACTCCATC 3262
QY 1373 CCGAAGAGTGACCGTGCAGCCCTCGAGTGCCTGCGGAGAGAGAGTGGACCGTGAACG 1432
Db 3263 CTGATAATGGACAGTACAGCCTATATGCTGCCAGAAAGACAGCTGGACTGTCAATG 3322
QY 1433 ACATCCAGAGCTGGTGGCAAGCTGAAGTGGCGCAGCAGATCTACCCCGCATCAAG 1492
Db 3323 ACATACAGAAAGTTAGTGGGAAATTTGAATGGCAAGTTCAGATTTATCGAGGATTAAG 3382
QY 1493 TGCCCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTGACCGACATCGTGGCCCCGA 1552
Db 3383 TAAAGCAGTTATGTAAACTCCTTAGAGGAAACCAAGCACTAACAGAAATATACCACTAA 3442
QY 1553 CCGAGGAGCGGAGTGGAGCTGGCCGAGAACCGCGAGATCCTCGCGAGCGCCCTGCAAG 1612
Db 3443 CAGNAGAGCAGAGCTAGACTGGCAGAAACACAGGAGATCTTAAAGAACCATGATCATG 3502
QY 1613 GCCTGTACTACGACCCCGCAGCAAGCCTGGTGGCGGAGATCCAGAAAGCGGCCACGACC 1672
Db 3503 AAGTATATTTAGACCATCAAAAGACTTTAGTAGCAGAAATACAGAAAGCAGGGCAAGGCC 3562
QY 1673 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTGGAAGACCGGCAAGTACGCCA 1732
Db 3563 AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACACAGAAAGTATGCA 3622
QY 1733 AGATGCCACCGCCCAACCAACGCTGAAGCAGCTGACCGAGCGGTGAGAGAAATCG 1792
Db 3623 GGATGAGGGGTGCCACACTAATGATGTAAACAGTTAACAGAGCAGCTGCAAAAGTAT 3682
QY 1793 CCATGGAGAGATCGTGATCTGGGCGAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1852
Db 3683 CCACAGAAAGCATAGTAAATATGGGAAAGATTCCTAAATTTAAACTACCCATACAAAG 3742
QY 1853 AGACTGGAGACCTGCTGGACCGACTACTGGCAGGCGCCCTGATCCCGAGTGGGAGT 1912
Db 3743 AAACATGGGAGAGCATGGTGGATGGAGTATTGGCAAGCTACTCTGGATTCTCTGAGTGGGAGT 3802
QY 1913 TCGTGAACACCCCGCTCGTGAAGCTGTGTACAGCTGGAGTGGAGAGGAGCCCATCATCG 1972
Db 3803 TTGTCAATACCCCTCCCTTAGTGAATTTAGTACCAGTTAGAGAAAGAACCCATAGTAG 3862
QY 1973 GCGCGAGACCTTCTAGTGGACCGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCG 2032
Db 3863 GAGCAGAAACTTTCTATGTAGATGGGAGCTAATAGGAGACTTAAATTAGGAAAGAGCAG 3922
QY 2033 GCTACGTGACCGACCGCGCGGCGCAGAAAGTCTGTAGCTGACCGAGACCAACACAGA 2092
Db 3923 GATATGTTACTGACAGAGGAGAACAAAAGTTGTTCCATAGCTGACACAAACATCAGA 3982
QY 2093 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCGAGGACAGCGGAGGAGTGAACATCG 2152
Db 3983 AGACTGAATTACAGCAATTCATCTAGCTTTGCGAGGATTCGGGATTAGAACTAAACATAG 4042
QY 2153 TGACCGCAGCAGTACGCCCTGGGCATCATCTCAGGCGCCAGCCCGACAGAGCAGAGCG 2212
Db 4043 TAACAGACTCACATATGCTATTAGTAATTCATTCAAGCACAAACACAGATAAGAGTGAATCAG 4102
QY 2213 AGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAGAGGAGAAAGTGTACCTGAGCTGG 2272
Db 4103 AGTTAGTCAAGTCAAAATAGAGCAGGTTAATAAAAAAGGAAAGGTTACCTGCGATGGG 4162

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	1632.6	66.3	4307	10	US-09-999-183-2	Sequence 2, Appli
2	1566	63.6	8366	10	US-09-872-733-6	Sequence 6, Appli
3	1530.6	62.1	4338	10	US-09-872-733-1	Sequence 1, Appli
4	1163.8	47.3	2467	10	US-09-872-733-3	Sequence 3, Appli
5	1106.8	44.9	2601	10	US-09-735-487-7	Sequence 7, Appli
6	1106.8	44.9	2601	10	US-09-735-487-9	Sequence 9, Appli
7	1106.8	44.9	2601	10	US-09-735-487-11	Sequence 11, Appli
8	1106.8	44.9	2601	10	US-09-735-487-13	Sequence 13, Appli
9	1106.8	44.9	4307	10	US-09-999-183-1	Sequence 1, Appli
10	1106.8	44.9	9719	10	US-09-737-190A-1	Sequence 1, Appli
11	1054.4	42.8	8933	10	US-09-943-286-3	Sequence 3, Appli
12	1054.4	42.8	8933	10	US-09-943-286-4	Sequence 4, Appli
13	1054.4	42.8	8933	10	US-09-943-286-9	Sequence 9, Appli
14	1048	42.5	9544	10	US-09-798-675-4	Sequence 4, Appli
15	1048	42.5	9918	10	US-09-798-675-5	Sequence 5, Appli
16	843.6	34.3	2348	9	US-10-097-997-1	Sequence 1, Appli
17	831	33.7	9793	9	US-09-886-156-56	Sequence 56, Appli
18	831	33.7	9793	9	US-09-886-150-56	Sequence 56, Appli
19	831	33.7	9793	9	US-09-886-149-56	Sequence 56, Appli

Db 1266 GATGAAGACTGTACTGAGAGACAGGCTAA-TTTTTTAGGAAGATCTGGCCCTTCTCTACA 1324
Qy 249 GGGCAAGGCCGGGAGTTCCCAAGCAGAGAACCGCGCCAAAGCCCAACAGCCCGACCGCCGA 308
Db 1325 AGGAAGGCCAGGGAATTTCTTCAGAGCAGAGACAGAGCCCAACAGCCCAACAGAGAGA 1384
Qy 309 GCTGCAAGTGGCGGG-----CGACAAACCCCGCAGAGCGCGCGCGCGCGCGCGCAGG 362
Db 1385 GCTTCAGTCTGGGGTAGACACAACACTCCCTCAGAAGCAGGAGCCGATAGACAGG 1444
Qy 363 CA-----CCCTGAACCTCCCGAGATCACCCCTGTGACGCGCCCGCTGTGTGAGATCAA 416
Db 1445 AACTGTATCTTTAACTTCCCTCAGATCACTCTTTGGCAACGACCCCTCTCTCAATAAA 1504
Qy 417 GTTGGGGGGCAGATCAAGAGAGCCCTGTGGACACCGCGCGCGCGCGCGCGCGCGCTGTGA 476
Db 1505 GATAGGGGGCAGCTCAAGAGAGCTCTCTGGACACCGGACAGACACACCGCTGTGGA 1564
Qy 477 GGAGATGAGCTGCGCCGCAAGTGGAAAGCCCAAGATGATCGGGGCAATCGCGCGCTTCAT 536
Db 1565 GGAGATGCTGTGCCAGCCGCTTGAAGCCGAGATGATCGGGGATCGCGGTTTCAT 1624
Qy 537 CAAGGTGGCCAGTACGACCAAGATCTGTATGAGATCTGCGGCAAGAGCCATCGCGAC 596
Db 1625 CAAGTGGCCAGTATGACCAAGATCTCTCATCGAAATCTGGGCGCACAAAGCTATCGGTAC 1684
Qy 597 CGTCTGATCGGCCCAACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGG 656
Db 1685 CGTCTGGTGGGCGCCACACCCGTCACATCATCGAGCAACCTGTGTGACGACATCGG 1744
Qy 657 CTGACCCCTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGTGAAGCTGAAGCCCGG 716
Db 1745 TTGCACGCTGAACCTTCCCATTAGCCCTATCGAGACGATACCGGTGAAGCTGAAGCCCGG 1804
Qy 717 CATGAGCGCCCAAGTGAAGCTGAGAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGAC 776
Db 1805 GATGAGCGCCGCAAGTCAAGCAATGCCATTCAGAGAGAGAGATCAAGGCACCTGGT 1864
Qy 777 CGCCATCTCGGAGGATGGAAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCC 836
Db 1865 GGAGATTGACAGAGATGGAAGAGGAGGAATCTCCAGATTTGGGCTGAGAACCC 1924
Qy 837 CTACAACACCCCGCTTTCGCGCATCAAGAGAGAGGACAGCACCAAGTGGGCGCAAGCTGGT 896
Db 1925 GTACAACACCGCGGTGTCGCAATCAAGAGAGAGGACTCTGACGAAATCGGCGCAAGCTGGT 1984
Qy 897 GGACTTCGGGAGTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCC 956
Db 1985 GGACTTCGGGAGTGAACAGCCACGCAAGACTTCTGGAGGTTTCAAGCTGGGCATCCC 2044
Qy 957 CCACCCCGCGGCTGAAGAGAGAGAGCGTGACCGTGTGACGCTGGGCGACGCTA 1016
Db 2045 GCACCCCGAGGCTGAAGAGAGAAATCCGTGACCGTACTGGATGGGTGATGCTA 2104
Qy 1017 CTTGAGGCTGGCCCTGGACAGGACTTCCGCAAGTACACCGCCCTTACCATTCGCCAGAT 1076
Db 2105 CTTCTCGGTTCCCTGGACGAAAGCTTCAGGAAGTACACTGCCCTTCAATTCCTCGAT 2164
Qy 1077 CAACAACGAGACCCCGCATCCGCTACCACTACAAGTGTGTCGCCCGCAGGCTGGAAGG 1136
Db 2165 CAACAACGAGACCCGGGATTCGATATGATACAGCTGTGTCGCCCGCAGGCTGGAAGG 2224
Qy 1137 CAGCCCGAGCATCTTCAGAGCAGCATGACCAAGATCTGTGAGCCCTTCCCGCCCGCCAA 1196
Db 2225 CTCCTCCGCAATCTCCAGAGTAGCATGACCAAAATCTGTGAGCCTTTCGCGCAACAGAA 2284
Qy 1197 CCGCGAGATCTGTATCTACCA-----GGCCCCCTGTAGCTGGGCGACCGACCTGGAGAT 1250
Db 2285 CCGCGAGATCTGTATCTATCATAGTACATGATGACTTGTACGTGGGCTCTGATCTAGAGAT 2344
Qy 1251 CGGCGAGACCCCGCAGATCGAGGAGCTGCGCAAGACACCTGCTGCGCTGGCGCTTAC 1310
Db 2345 AGGCGAGCCCGCACCAAGATCGAGGAGCTGCGCCAGCACCTGTGTGAGTGGGACTGAC 2404

Qy 1311 CACCCCGCACAAGACCAAGAGAGAGCCCTTCTCTGTGATGGCTTACGAGCTGCA 1370
Db 2405 CACACCCGACAAAGACCAAGAGAGCCCTTCTCTGTGATGGTTACGAGCTGCA 2464
Qy 1371 CCCGCAAGTGGACCGTGCAGCCCATCGAGTGCCTCGAGAGAGAGCTGGACCGTGAA 1430
Db 2465 CCCTGACAAATGGACCGTGCAGCTTATCGTGTGCCAGAGAAAGACAGCTGGAGTGTCAA 2524
Qy 1431 CGACATCCAGAAGTGTGGCAGAGTGAATGGCCAGCCAGATCTACCCCGCATCAA 1490
Db 2525 CGACATACAGAAGCTGGTGGGGAAGTTGAATGGCCAGTCAAGATTTACCCAGGATTTAA 2584
Qy 1491 GGTGCGCAGCTGTGCAAGCTGTGCGCGCGCAAGGCCCTGACCCAGATCTGTCGCCCT 1550
Db 2585 GTGTGAGCAGCTGTGCAAACTCTCCCGGACCAAGGACTACAGAGGTGATCCCGCT 2644
Qy 1551 GACCGAGAGCGGAGCTGGAGTGGCGGAGAACCGGAGATCTGCGGAGCCCGTGCA 1610
Db 2645 AACCGAGAGCGGAGCTCGAACTGGCAGAAACCGAGAGATCTTAAGGAGCGCGTGCA 2704
Qy 1611 CGCGTGTACTACGACCCAGCAGGACCTGGTGGCGGAGATCCAGAGCAGGCGCCACGA 1670
Db 2705 CGCGTGTACTATGACCCCTCCAAAGGACCTGATCGCCGAGATCCAGAGCAGGCGCAGG 2764
Qy 1671 CCAGTGGACCTTACCAGATCTACCAGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGC 1730
Db 2765 CCAGTGGACCTTATCAGATTTACCAGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGC 2824
Qy 1731 CAGATCGGACCCCGCACACCAAGCAGAGTGAAGCAGTGAAGCAGGCGCGTGCAAGAT 1790
Db 2825 CCGATGAGGCGTCCCAACACATAACGACGTAACGACGTAACGAGGCGCGTGCAAGAT 2884
Qy 1791 CGCCATGAGAGCATCTGTGATCTGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAA 1850
Db 2885 CACCACCAAGACATCTGTGATCTGGGAAGAGCTCTTAAGTTCAAGCTGCCATCCAGAA 2944
Qy 1851 GGAGACCTGGGAGACCTGGTGACCGACTACTGCGAGGCCACTTGGATTCGCCGAGTGGGA 1910
Db 2945 GGAAACCTGGGAAACCTGGTGACAGAGTATTGGCAGGCCACCTGGATTCCTGAGTGGGA 3004
Qy 1911 GTTGTGAACACCCCGCCCTTGGTGAAGCTGTGTACAGCTGGAGGAGGAGCCCATCAT 1970
Db 3005 GTTGTGAACACCCCTTCCCTGGTGAAGCTGTGTACAGCTGGAGAGGAGGAGCCCATAGT 3064
Qy 1971 CGGCGCGAGACCTTCTACGTGGAGCGCGCCCAACCGCAGACCAAGATCGGCAAGGC 2030
Db 3065 GGGCGCGGAAACCTTCTACGTGGATGGGCGCGCTTAACAGGAGACTAAGCTGGGCAAGC 3124
Qy 2031 GGGTACGTGACCAACCGGCGCGGAGAGATCTGTGAGCCTGACCGAGACCAACCA 2090
Db 3125 CGGATACGTCACTAACCGGCGAGACAGAGTGTTCACCTCACTGACACCAACCA 3184
Qy 2091 GAACACGAGCTGACGCCATCCAGCTGGCCCTGCAGACAGCGGACCGAGGTGAACAT 2150
Db 3185 GAAGAGTGTGAGCGGCGCATTTTACCTGCTTGTGAGGACTCGGCGCTGGAGGTGAACAT 3244
Qy 2151 CGTACCGACAGCAGCTACGCCCTGGGCGCATCATCCAGGCCCGGCGGCAAGAGCGAGAG 2210
Db 3245 CGTGACAGCTCTCAGTATGCCCTGGGCGATCATTAAGCCCGCAGCAGACAGGTGATC 3304
Qy 2211 CGAGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTG 2270
Db 3305 CGAGCTGTGCTAATCAGATCATCGAGCAGCTGATCAAGAGGAAAGGCTATCTATCTGCGCTG 3364
Qy 2271 GGTCCCGCCCAAGGCGCATCGGCGCAACGACAGATCGACAGCTGGTGGAGCAAGG 2330
Db 3365 GGTACCCCGCCCAAGGCGATTTGGGCGAATGAGCAGGTGCAAGAGCTGCTCGGCTGG 3424
Qy 2331 CATCCGCAAGTGTCTTCCTGGAGCGCATCGA 2363
Db 3425 CATCAGGAAGTGTCTATTCCTGGATGGCATCGA 3457

RESULT 2

```

US-09-872-733-6
; Sequence 6, Application US/09872733
; Patent No. US2001003655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, HIV GAG AND
; TITLE OF INVENTION: HIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, HIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBkan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733-6

```

Query Match	63.6%;	Score 1566;	DB 10;	Length 8366;
Best Local Similarity	80.3%;	Pred. NO. 5.6e-252;		
Matches 1905; Conservative	0;	Mismatches 445;	Indels 22;	Gaps 5;

QY	14	TGCCCCAGGCCATGAGCCAGGCCACCAGG---GCCAACATCCTGATGAGCGCAGCAACT	70
-			
Db	1857	TGCCCCAGGCCGATGAGCCAGGAGTGAGCGAACTCGGGGACCATAATGATGAGAGGAGCACT	1916
QY	71	TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCACTGCGGCAAGAGGGCCACATCGCCC	130
Db	1917		
QY	131	GCAACTGCGCGCCCCCGCCAGAAAGGCTGCTGGAAGTGCAGCAAGAGGGCCACCAGA	1976
Db	1977	GGAACGTGCGGGCCCCCGGAAGAAGGGCTGTGGAAATGTGGAAAGGAAGGACACCCAAA	2036
QY	191	TGAAGGACTGCACCGAGCGCCAGGCCACACTTCTCCGGCAGGACCTGGCCCTTCCCCAGG	250
Db	2037		
QY	251	GCAAGCCCGCGAGTTCCCGAGCAGAGAACCGCGGCCAACAGCCCCACACAGCGCGGAGC	310
Db	2096	GGAGGCCAGGGAATTTCTTCAGAGCAGACACAGGCCACAGCCCCACAGAAAGAGGC	2155
QY	311	TGCAGTGGCGGG-----CGAACACCCCGACGAGGCGCGCGCGCGAGCGCCAGGCA	364
Db	2156	TTTCAGGTCTGGGGTAGAGACACAACTCCCCCTCAGAAGCAGGAGCGATAGACAAGGAA	2215
QY	365	-----CCTGTGAATTCCTCCAGATCACCTGTGGCAGCGCCCTGGTGAGCATCAAGG	418
Db	2216	CTGTATCTTTTAATTTCCCTCAGATCACTTTTGGCAACGACCCCTCGTCACAGTAGGA	2275
QY	419	TGGCGCGCCAGATCAAGAGGCGCTGTGTGACACCGCGCGCGCGAGCACCGTCTGGAGG	478
Db	2276	TCGGGGGCACTCAAGNAGCGCTGCTCGATACAGAGCAGATGATACAGTATTGAAG	2335
QY	479	AGATGAGCTGCCGGCAAGTGGAGCCCAAGATGATCGCGGCGATCGCGGGCTTCATCA	538
Db	2336	AAATGAGTTTGCAGGAAGATGGAAACCAAAAATGATAGGGGGGATCGGGGGCTTCATCA	2395
QY	539	AGGTGGCCAGTACACACAGATCCTGATCGAGATCTCGCGCAAGAGGCCATCGGCACCG	598
Db	2396	AGGTGAGGCGGTACACACAGATCTCTAGAAATCTCTGGACATAAAGCTATAGGTACAG	2455

Db	3536	ATGGACCTACCAATCTACGAGGACCTTCAAGACCTGAAGACAGGCAAGTACGCAA	3595
QY	1733	AGATGGCGCACCGCCACACACACAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG	1792
Db	3596	GGATGAGGGTGCCACACACACACAGTGAAGCAGCTGCAGAGGCAAGTGCAGAAATCA	3655
QY	1793	CCATGAGAGCATCGTATCTGGGCAAGACCCCAAGTTCCGCGCTGCCATCCAGAAGG	1852
Db	3656	CCACAGAGCATCGTATCTGGGCAAGACTCCCAAGTTCAAGCTGCCCATACAGAAGG	3715
QY	1853	AGACTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTTGATCCCGAGTGGGAGT	1912
Db	3716	AGACATGGGAGACATGGTGGACCGAGTACTGGCAGGCCACCTTGATCCCTGAGTGGGAGT	3775
QY	1913	TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGCTGGCAAGAGCCCATCATCG	1972
Db	3776	TCGTGAACACCCCTCCCTTGGTGAACCTGTGTATCAGCTGGCAAGAACCCATCGTGG	3835
QY	1973	GGCCCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGG	2032
Db	3836	GAGCAGAGACCTTCTACGTGGATGGGCGAGCCAAACAGGGAGACCAAGCTGGGCAAGGCG	3895
QY	2033	GCTACGTGACCGACCGGGCCGGCAGAGAGATGCTGAGCGTACCGAGACACCAACCCAGA	2092
Db	3896	GCTACGTGACCAACCGAGGAGACAGAAAGTGGTGACCTGACTGACCAACCAACCCAGA	3955
QY	2093	AGACCGAGCTGACAGGCCATTCAGCTGGCGCTGCAGGACAGCGGCGAGGTGAACATCG	2152
Db	3956	AGACTGAGCTGCAGCCATCTACCTAGCTCTGCAGACAGCGGACTGGAAGTGAACATCG	4015
QY	2153	TGACGACAGCCAGTACGCCCTGGGCATCATCCAGGCCACGCCGACAGAGCGAGAGCG	2212
Db	4016	TGACAGACTCACAGTACGCACTGGGCGATCTCCAAAGCACAAACCCAGACCAATCCGAGTCA	4075
QY	2213	AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAAGGAAAGGTGTACCTGAGCTGGG	2272
Db	4076	AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAAGGAAAGGTGTACCTGAGCTGGG	4135
QY	2273	TGCCCCGCCAAGGGCATCGCGGCACACGAGCAGATCGCAAGCTGGTGAGCAAGGCA	2332
Db	4136	TACCAACACAAAGGAATGGAGGAAATGAACAGTAGATAAATAGTCAGTCCCTGGGA	4195
QY	2333	TCGCGAAGTGTGTTTCCTGGACGGCATCGAT	2364
Db	4196	TCGCGAAGTGTGTTTCCTGGACGGGATCGAT	4227

```

RESULT 3
US-09-872-733-1
; Sequence 1, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287U51 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733-1

```

Query Match				62.1%	Score 1530.6;	DB 10;	Length 4338;
Best Local Similarity				80.1%	Pred. No. 4e-246;		
Matches 1905;				Conservative	0;	Mismatches 444;	Indels 28; Gaps
Qy	14	TGCGCGAGGCATGAGCCAGGCGACACAG--	--GCCAACATCCTGATGCAGCGAGCAACT	70			
Db							
Db	1085	TGCGCGAGGCGATGACCCAGGTGACAACT	CGGCGACCATTAATGATGCAGAGGCAACT	1144			
Qy	71	TCAAGGCCCCCAAGCCCATCATCAAGTGCTTCAACT	TCGGCAAGGAGGGCCACATGCGCC	130			
Db	1145	TCGGAAACACGCGAAGATCGTCAAGTGCTTCAATT	TGTGGCAAGAGGCGACACCGCCA	1204			
Qy	131	GCAACTGCGCGCCCCCGCCAGAAAGGCTGCTGGAAGT	TCGGCAAGGAGGGCCACAGGCGCCACCA	190			
Db	1205	GGAACATGCGGGCCCCCGGAAAGAGGCTGTGTGAAAT	GTTGGAAGATGTGGAAGAGGAGGACACCAA	1264			
Qy	191	TGAAGGACTCACCGAGCGCCAGGCCAATCTTCCGCGAGGAGCT	TGGCCTTCCGCCAGG	250			
Db	1265	TGAAGATTGTAATGAGACAGGCTAA--TTT	TTTATAGGAAGATCTGGCCTTCTCTACAAG	1323			
Qy	251	GCAAGGCCCGCGATTTCCCGAGGACGAAACCGCGCCCA	CAGGCCCCACAGCGCGCAGC	310			
Db	1324	GGAAGGCCAGGGAATTTCTTCAGAGCAGACACAGAGCC	CAACAGCCCCACCAAGAGAGC	1383			
Qy	311	TGCAGGTGCGCGG-----CGACACCCCGCAGCAG	AGCGCGCGCGCAGCGCCAGGCGCA	364			
Db	1384	TTCAGGTCTGGGTAGAGACACAACTCCCCCTCAG	AGCAGGACCGGATAGACAGAA	1443			
Qy	365	-----CCCTGAACTTCCCCAGATCACCCCTGTG	CAGCGCCCTGTGTGAGATCAAG	418			
Db	1444	CTGTATCCTTTAACTTCCCTCAGATCACATCTTT	TGGCAACGCCCTGTCTACAGTAAGA	1503			
Qy	419	TGGCGCCAGATCAAGGAGGCCCTGCTGGACACCG	CGCGCAGCAGCACCGTGTGGAGG	478			
Db	1504	TCGGGGGCAACTCAAGGAAGCGCTCTCGATAC	AGGACAGATGTATAGATTTAGAG	1563			
Qy	479	AGATGAGCCTTCCCGGCAAGTGGAAAGCCCAAGAT	GATCGCGGCATCGCGGCTTCATCA	538			
Db	1564	AAATGAGTTTGCCAGGAAGATGAAACCAAAAT	GATAGGGGGATCGGGGCTTCATCA	1623			
Qy	539	AGGTGCGCAGTACGACCAGATCCTGATCGATCT	CGGCAGAGAGCCATCGGCACCG	598			
Db	1624	AGGTGAGGCGAGTACGACCAGATACTCATAGAAT	CTGTGGACATTAAGCTATAGGTACAG	1683			
Qy	599	TGCTGATCGGCCCCAC-----CCCGTGAACAT	CATCGCGCGCAACATCTGACCCAGCTG	654			
Db	1684	TATTAGTAGACCTACCTACACCTGTACATTAATT	GGAAGAAATCTGTTGACCCAGATC	1743			
Qy	655	GGCTGCACCTGAACTTCCCATATCAGCCCATCG	AGACCGTGCCTCGTGAAGCTGAGGCC	714			
Db	1744	GGCTGCACTTGAACCTTCCCATATCAGCCCTAT	TGTAGAGCGTGCCTGTGAAGCTTGAAGCCG	1803			
Qy	715	GGCATGACGSCCCCAAGGTGAGCAGTGCGCCCT	TGACCGGAGGAAGATCAAGGCCCTG	774			
Db	1804	GGATGAGCGGCCCAAGGTCAAGCAATGGCCAT	TGACGAAGAGAGATCAAGGCTTA	1863			
Qy	775	ACCGCCATCTCGGAGGAGATGGAGAAGGAGG	CAAGATCAACCGCCCCGAGAAC	834			
Db	1864	GTCGAAATCTGTACAGATGGAGAAGGAAG	ATCAGCAAGATCGGCGCTGAGAAC	1923			
Qy	835	CCCTACAAACCCCGTGTTCGCCATCAAGAGA	AGGACAGCAACCAAGTGGCGCAGCTG	894			
Db	1924	CCCTACAAACTCCAGTCTTTCGCAATCAAGA	AGAAGGACGATPACCAAGTGGAGAAGCTG	1983			
Qy	895	GTGACTTCCGCGAGCTGAAACAGCGCACCC	CAGGACTTCTGGAGGTTGAGCTGGGCATC	954			
Db	1984	GTGACTTCAGAGCTGAAACAGAGAACTCAG	GAATCTTGGAAATTCAGCTGGGCATC	2043			
Qy	955	CCCCACCCCGCGCCTGGAAGAAGAAGAGCG	GTGACCGTCTGAGCTGGGCGACGCC	1014			
Db	2044	CCACATCCCGTGGTGTGAAGAAGAAGTCA	GTTGAGTCTGGATGGGTGATGCC	2103			

QY 1015 TACTTACGCGTCCCGTGGAGAGGACTTCGCAAGTACACCGCTTACACATCCCGAGC 1074
Db 2104 TACTTCCGCTTCCCTTGGAGGAGACTTCAGGAAGTACACTGCTTTCAGATACCTAGC 2163
QY 1075 ATCAACAACAGAGACCCCGGATCCCGTACAGTACAACTGCTGCGCCAGGCTGGAAG 1134
Db 2164 ATCAACAACAGAGACCCAGGATCCCGTACAGTACAACTGCTGCGCCAGGATGGAAG 2223
QY 1135 GGCAGCCCGAGCATTTCCAGAGCAGCATGACCAAGATCTGAGAGCCCTTCCGCGCCCGC 1194
Db 2224 GGATCACCGACCATCTTTCAAGCAGCATGACCAAGATCTTGGAGCCCTTCCGCAAGCAA 2283
QY 1195 AACCCCGAGATCGTGTATCTACCA-----GGCCCCCTGTACFTGGGCGAGCACTGGAG 1248
Db 2284 AACCCAGACATCGTGTATCTATCAGTACATGGAACCTCTAGGTAGGAAGTGTACCTGGAG 2343
QY 1249 ATC-GGGCAGCAGCGCCGCAAGATCGAGGAGCTGCGCAAGCACTGCTGCGCTGGGCGTT 1307
Db 2344 ATCGGGCAGCAGCAGGACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGGTGGGACT 2403
QY 1308 CACCACCCCGACAAAGAACACCAAGAGGAGCCCTCTCTGTGGATGGGCTACGAGCT 1367
Db 2404 GACCACACCAGCAAGAACGACCAAGAGGAGCACTTCTCTGTGGATGGGCTACGAGCT 2463
QY 1368 GCACCCCGACAAGTGGACCTGACGCCATCGAGCTGCCGAGAGGAGAGCTGGACCGT 1427
Db 2464 GCATCTTGACAAGTGGAGAGCTGACGCCATCGTCTGCTGAGAGGAGACAGCTGAGCTGT 2523
QY 1428 GAAGGACATCCAGAACTGGTGGGCAAGCTGAAGTGGGCGAGCCAGATCTACCCCGGAT 1487
Db 2524 GAAGGACATACAGAACTGGTGGGCAAGCTGAAGTGGGCGAGCCAGATCTACCCAGGAT 2583
QY 1488 CAAGTGGCGCAGCTGTGAAGCTGCTGCGGCGGCAAGGCGCTGACCGACATCTGTGCC 1547
Db 2584 CAAGTGGCGCAGCTGTGAAGCTGCTGCGGCGGCAAGGCGCTGACCGACATCTGTGCC 2643
QY 1548 CCTGACCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGCGAGCCGT 1607
Db 2644 ACTGACAGAGGAGGAGCTGAGAACTGGCAGAGAACCGGAGATCTCTGAGGAGCCAGT 2703
QY 1608 GCAGGCGGTACTACGACCCCGACAGGAGCTGGTGGCGGAGATCCAGAGGAGGCGCA 1667
Db 2704 ACATGGAGTGTACTACGACCCAGCAAGGAGCTGATCGCAGAGATCCAGAGGAGGCGCA 2763
QY 1668 CGACCATGGACCTTACAGATCTACAGGAGCCCTTCAAGAACTGAAGCCGCAAGTA 1727
Db 2764 AGGCAATGACCTTACCAATCTACAGGAGCCCTTCAAGAACTGAAGCCGCAAGTA 2823
QY 1728 CGCAAGATGCGCACCCCGCACACACGAGCTGAAGCAGCTGACCGAGGCGCTGCGAGAA 1787
Db 2824 CGCAAGATGAGGCGTCCCGCACACACGAGTGTGAAGCAGCTGACAGAGGAGTGCAGAA 2883
QY 1788 GATCGCATGGAGAGCATGTGATCTGGGCAAGACCCCGCAAGTTCGCGCTGCCATCCCA 1847
Db 2884 GATCACCACAGAGAGCATGTGATCTGGGCAAGACTCCCAAGTTCAGAGCTGCCATACA 2943
QY 1848 GAAGGAGACCTTGGAGACCTTGGTGGACCTACTGCGAGCCACCTGGATCCCGGAGTG 1907
Db 2944 GAAGGAGACATGGAGAGCATGTGGGACCGAGTACTGGCAAGCCACTGGATCCCTGAGTG 3003
QY 1908 GGAGTTCGTGAACACCCCGCTGGTGAAGCTGTGGTACAGCTGGAGAGGAGGAGCCAT 1967
Db 3004 GGAGTTCGTGAACACCCCTCCCTTGGTGAAGCTGTGGTATCAGCTGGAGAGGAGCCAT 3063
QY 1968 CATCGGCGGAGACCTTCTACGTGGAGGCGCGCGCAACCGCGAGACCAAGATCGGCA 2027
Db 3064 CGTGGGAGCAGAGACCTTCTACGTGGATGGGCGAGCCCAAGGAGAGCAAGCTGGGCA 3123
QY 2028 GCGCGGTACGTGACCGGCGCGGAGAGATCTGTAGCTGTGAGCCGTGACCGAGACCCAA 2087
Db 3124 GCGAGGTACGTGACCGGCGCGGAGAGCAAGAGTGTGAGCCCTGATGACACCCAA 3183
QY 2088 CCAGAGACCGAGCTGCAAGGCGCATCCAGCTGGCCCTGCAAGGAGGCGGAGGAGTGA 2147

Db 3184 CCAGAGACTGAGCTGCAAGCCATCTACCTAGCTCTGCAAGACAGGAGCTGGAAGTCAA 3243
QY 2148 CATCGTACCGACAGCAGTACGAGCCCTGGGATCATCCAGGCCCGCCGACAGAGCGGA 2207
Db 3244 CATCGTACAGACTTCACAGTACG-CATGGGATCATCCAGCAACACAGACCAATCGGA 3302
QY 2208 GAGGAGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAG 2267
Db 3303 GTCAGAGCTGCTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGCGC 3362
QY 2268 CTGGTCCCGCCCGACAGGCGATCGGCGCAACGAGCAGATCGACAAAGCTGGTGAGCAA 2327
Db 3363 ATGGGTACCAACACACAAAGGAATGGAGGAAATGAACAAGTAGATAAATTAGTCAGTGC 3422
QY 2328 GGGCATCCGCAAGGTGCTGTTCTTCTGGAGCGCATCGAT 2364
Db 3423 TGGATCCGGAAGGTGCTGTTCTTCTGGAGCGGATCGAT 3459

RESULT 4
US-09-872-733-3
; Sequence 3, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287051 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
US-09-872-733-3

Query Match 47.3%; Score 1163.8; DB 10; Length 2467;
Best Local Similarity 84.0%; Pred. No. 3e-185;
Matches 1328; Conservative 0; Mismatches 247; Indels 6; Gaps 1;
QY 790 GAGATGGAGAGGAGGCGCAAGATCACCAGATCGGCCCGGAGAACCCCTACACACCCGCC 849
Db 7 GAGATGGAGAGGAGGAGGAGATCAGCAAGATCGGCCCTGAGAACCCCTACACACTCCA 66
QY 850 GTGTTGCGCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGGACTTCGCGAG 909
Db 67 GTCITCGCAATCAAGAAGAAGACAGTACCAGTGGAGAGAGCTGGTGGACTTCAGAGAG 126
QY 910 CTGAACAGCCAGCCAGGACTTCTGGAGGTGCAGCTGGGCGATCCGCCAGCCCGCGGC 969
Db 127 CTGAACAAGAGAACTCAGGACTTCTGGGAAGTTCAGTGGGATCCCATCCAGCTCCGCTGG 186
QY 970 CTGAAGAAGAAGAAGCGCTGACCGCTGCGAGCTGGGCGAGCGCTACTTTCAGCTGGCC 1029
Db 187 TTGAAGAAGAAGAAGTCAAGTACAGTGTGATGCTGGATGCTGGTGTGCTTCTCCGTTCCC 246
QY 1030 CTGGACGAGGACTTCCGCAAGTACACCCCTTACCATCCCGAGCATCAACACAGAGACC 1089
Db 247 TTGACGAGGACTTCAGGAAGTACACTGCTTCCAGTACCTAGCATCAACACAGAGACA 306
QY 1090 CCCGGCATCCGCTACCAAGTACAACTGCTGCCCGCAGGCTGGAAGGAGGCCGCCAGCATC 1149
Db 307 CCAGGATCCGCTACCAAGTACAACTGCTGCCACAGGATGGAAGGATGGAAGGATCACCAGCCATC 366

QY 1150 TTCAGAGCAGCATGACCAAGATCCTTGAGCCCTTCCGGCCGCCAACCOCAGAGATCGTG 1209
Db 367 TTTCAAGCAGCATGACCAAGATCCTTGAGCCCTTCCGCAAGCAAAACCCAGACATCGTG 426
QY 1210 ATCTACCA-----GGCCCCCTGTAGTGGGAGAGACCTGGAGATCGGCCAGCACCGC 1263
Db 427 ATCTATCAGTACATGAGGAGGACCTTACGTAGGAAGTACCTGGAGATCGGGCAGCAGG 486
QY 1264 GCCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCTGGGGCTTCACACCCCGACAAG 1323
Db 487 ACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGTGGGGACTGACCACACAGACAAG 546
QY 1324 AAGCACCAGAGAGAGCCCCCTTCTGTGATGGCTTACGAGCTGCACCCCGACAAGTGG 1383
Db 547 AAGCACCAGAGAGAACCTCCCTTCTGTGATGGCTTACGAACCTGCATCTGCAACAAGTG 606
QY 1384 ACCGTGAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGAGCGTGAACAGATCCAGAAG 1443
Db 607 ACAGTGCAGCCCATCGTGTGCTGGAAGAGACAGCTGGAGCTGAACGACATACAGAAG 666
QY 1444 CTGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAGTGGCGCAGCTG 1503
Db 667 CTGTGGGCAAGTGAACCTGGGCAAGCAGATCTACCCAGGCATCAAGTTAGCAGCTG 726
QY 1504 TGCAGCTGTGCGCGGCCAAGGCCCTGACCGACATCTGTGCCCTGACCGAGAGGCC 1563
Db 727 TGCAGCTGTGCGCGGCCAAGGCCACTGACAGAGTGTCCACTGACAGAGGAACCA 786
QY 1564 GAGCTGAGCTGCGCGAGAACCCGAGATCTTGGCGAGCCCTGACCGCGGTGACTAC 1623
Db 787 GAGCTGAGCTGCGCGAGAACCCGAGATCTTGAAGGAGCCATACATGGAGTGTACTAC 846
QY 1624 GACCCAGCAGGAGCTGTGGCGGAGATCCAGAAGCAGGGCCACGACCTGAGACCTAC 1683
Db 847 GACCCAGCAGGAGCTGTGCGAGATCCAGAAGCAGGGCCAGGCCAATGGACCTAC 906
QY 1684 CAGATCTACGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATCGCAC 1743
Db 907 CAATCTACGAGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGAGGGGT 966
QY 1744 GCCACACCAAGCAGTGAAGCAGCTGACCGAGGCGCTGAGAGATCGGCATGGAGAGC 1803
Db 967 GCCACACCAAGCAGTGAAGCAGCTGACAGAGCAGTGCAGAGATTCACCAAGAGAGC 1026
QY 1804 ATCGTATCTGGGCAAGACCCCAAGTTCCGCTGCGCCATCCAGAGAGACCTGGGAG 1863
Db 1027 ATCGTATCTGGGCAAGACTCCCAAGTTCAAGCTGCCATACAGAGAGAGATGGGAG 1086
QY 1864 ACCTGTGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCTGTGAACCC 1923
Db 1087 ACATGTGGACCGAGTACTGGCAAGCCACCTGGATCCCTGAGTGGGAGTTCTGTGAACCC 1146
QY 1924 CCCCCCTGTGTAAGCTGTGTTACAGCTGAGAGAGAGCCCATCATCTGCGCGCGAGACC 1983
Db 1147 CTTCCCTTGTGTAAGCTGTGTTACAGCTGAGAGAGAGCCCATCTGTTGGGAGCAGAGCC 1206
QY 1984 TTCTACTGTGACGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGCTACGTGACC 2043
Db 1207 TTCTACTGTGATGGGCGACCAACAGGAGACCAAGCTGGCAAGCGCAGCTACGTGACC 1266
QY 2044 GACCGGGCGCGCAGAGATCTGTGAGCTGACCGAGACCAACCAACAGAGACCGAGCTG 2103
Db 1267 AACCGAGGACGACAGAAGTGTGACCCCTGACTGACACCAACCAACAGAGACTGAGCTG 1326
QY 2104 CAGGCATCCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCTGTCGACCGAGC 2163
Db 1327 CAGGCATCTACCTAGCTCTGCAAGACAGCGGAGTGAAGTGAACATCTGTCGAGACTCA 1386
QY 2164 CAGTACCCCTGGGCAATCTACAGGCCCGAGCCGCAAGAGCAGGAGCGAGCTGGTGAAC 2223
Db 1387 CAGTACGCACTGGGCAATCTACAGGACCAACAGACCAATCCCGAGTCTAGAGCTGGTGAAC 1446

QY 2224 CAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGGTGTACTGTGAGTGGTCCCGCCAC 2283
Db 1447 CAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACTGTGGATGGGTACCAGACAC 1506
QY 2284 AAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGACAAGGGCATCCGCAAGGTG 2343
Db 1507 AAAGGAATTGGAGAAATGAACAAGTAGATAAATTAGTCAAGTGTGGATCCGGAAGGTG 1566
QY 2344 CTGTTCTCTGGACGGCATCGAT 2364
Db 1567 CTGTTCTCTGGACGGATCGAT 1587
RESULT 5
US-09-735-487-7
; Sequence 7, Application US/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-735-487-7

Query Match 44.9%; Score 1106.8; DB 10; Length 2601;
Best Local Similarity 68.2%; Pred. No. 8.7e-172;
Matches 1618; Conservative 0; Mismatches 732; Indels 22; Gaps 5;

QY 14 TGGCCGAGGCGCATGAGCAGCGCCACCA--CGGCCAATCTCTGATGCGGCGAGCAACT 70
Db 77 TGGCTGAAGCAATGAGCCAAGTACAAATTCAGTCCATATGATGAGAGGCAATT 136
QY 71 TCAGGGCCCCCAAGCGGCATCATCAAGTCTTCAACTCGGCAAGGAGGGCCACATCGCC 130
Db 137 TTAGGAACCAAGAAAGATTGTTAAGTGTTCATTTCTGCAAAAGAGGGCACACGCCA 196
QY 131 GCAATGTCGCGCGCCCCCGCAAGAGGGCTGCTGGAGTGGGCAAGAGGGCCACACAGA 190
Db 197 GAATTCAGGGGCCCCCTAGGAAAAGGGCTGTTGAAATGTGAAAGAGGAGGACACAAA 256
QY 191 TGAAGGACTGCACGAGCGCGCAGGCCAACTTCTTCGCGAGGACCTGGCTTCCCGCAGG 250
Db 257 TGAAGATTGTTACTGAGAGACAGGCTAA-TTTTATAGGAGATCTTGGCTTCTTACAAG 315
QY 251 GCAAGGCCCGCGAGTTCCCGAGCGAGCAACCGCGCAACAGCCCCACCGCGCGAGC 310
Db 316 GGAAGGCCAGGGAATTTCTTCAGAGCAGACCAAGCCACAGGCCCCACCAAGAGAGAGC 375
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGGAGCGCGCGCGAGCGCGAGGCA 364
Db 376 TTCAGGTCTGGGCTAGAGACAACAATCTCCCTCAGAAAGCAGGAGCGGATAGACAAGAA 435
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCTCTGCTGAGCATCAAG 418
Db 436 CTGTATCCTTTAACTTCCCTCAGTCTACTCTTTGGCAACGACCCCTCTCACAATAAGA 495

QY 419 TGGGGGCGCAGATCAAGAGAGCCCTGTGGACACCGCGCCGAGACACCCGTGTGGAGG 478
Db 496 TAGGGGGGCACTAAAGAGAGCTCTATTAGATACAGGAGCAGATACACTATTAGAAG 555
QY 479 AGATGAGCCTCGCGGCAAGTGGAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 538
Db 556 AAATGAGTTTGGCAGGAAGATGGAAACCAAAATGATAGGGGGAATTGGAGGTTTATCA 615
QY 539 AGGTGCGCAGTAGCACACAGATCTGTGAGATCTGGCGCAAGAGCCATCGGCACCG 598
Db 616 AAGTAAGACAGTATGATCAGATACTCATGAATCTGTGGACATAAAGCTATAGGTACAG 675
QY 599 TGCTGATGGGCCCCACCCCGTGAACATCATCGGCGCAACATCGTACCCAGCTGGGCT 658
Db 676 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGTGACTCAGATTGGTT 735
QY 659 GCACCTGAACCTTCCCATACAGCCCATCGAGACGGTGGCCGTGAAGCTGAAGCCCGCA 718
Db 736 GCACCTTAAATTTCCCATATAGCCCTATTGAGACTGTACCAAGTAAATTTAAGCCAGGAA 795
QY 719 TGGACGGCCCAAGTGAAGCAGTGGCCCTGACCGGAGGAAGATCAAGCCCTGACCG 778
Db 796 TGGATGGCCCAAGTTAAACAATGGCCATTGACAGAGAAAATAAAGCATTAGTAG 855
QY 779 CCATCTGCGAGGAGTAGAGAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 856 AAATTGTGACAGAGTGAAGAGGAGGCAAAATTTCAAAAATTTGGGCTGAAATCCAT 915
QY 839 ACAACACCCCGTGTCCCATCAAGAAAGAGCAGCACCAAGTGGCGGAGCTGGTGG 898
Db 916 ACAATACTCCAGTATTGGCCATAAGAAAAGAACAGTACTGATGATGATGATGATGAT 975
QY 899 ACTTCGCGAGCTGAACAGCGCCACAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCC 958
Db 976 ATTCAGAGACTTAATAGAGACTCAAGACTTCTGGGAAGTTCAATTAGAAATACCAC 1035
QY 959 ACCCGCGCGGCTGAAGAAAGAGAGCGTGAACCGTGTGGAGCTGGGCGACGCTACT 1018
Db 1036 ATCCCGAGGCTTAAAAAGAAAAATCAGTAACAGTACTGGATGATGATGATGATGAT 1095
QY 1019 TCAGGCTGCCCTGGACGAGACTTCCCAAGTACACCGGCTTACCATCCCGACATCA 1078
Db 1096 TTTCAGTTCCCTTAGATGAAGACTTCAGGAAGTATCTGCAATTTACCATCTAGTATAA 1155
QY 1079 ACAACGAGACCCCGGCTCCGCTACCAAGTACAAGTCTGCCCGAGGCTGGAAGGCA 1138
Db 1156 ACATGAGACACAGGATTAAGATATAGTACATGCTTCCACAGGGATGGAAGGAT 1215
QY 1139 GCGCCAGCATCTCCAGAGCAGTACCAAGATCTCGAGCCCTTCCGCGCCGCAACC 1198
Db 1216 CACGAGCAATATCCAAAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAATC 1275
QY 1199 CCGAGATCGTATACCA-----GGCCCCCTGTAGTGGGAGCAGCAGCTGGAGATCG 1252
Db 1276 CAGACATAGTTATCTATCAATACATGATGATGATGATGATGATGATGATGATGATGAT 1335
QY 1253 GCCAGACCGCGCAAGATCGAGGAGCTGCGCAAGCACTCTCGCTGGGGCTTCACCA 1312
Db 1336 GCGAGCATAGACAAAAATAGAGAGCTGACACACATCTCTTGAGGTGGGACTTACCA 1395
QY 1313 CCCCCGACAGAACACAGAGAGGAGGCCCCCTTCTGTGGATGGGCTACGAGCTGACC 1372
Db 1396 CACCAGAAAAAACATCAGAAAGAACCTCCATTTCTTGTGATGGTTATGAACCTCCATC 1455
QY 1373 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCGCGAGAGAGGAGCTGGACGTGAACG 1432
Db 1456 CTGATAATGACAGTAGTACAGCTTAGTGTGCGCAAGAAAGACAGCTGGACTGTCAATG 1515
QY 1433 ACATCCAGAACTGGTGGGCAAGCTGAACCTGGGCGAGCAGCATCTACCCCGCATCAGG 1492
Db 1516 ACATACAGAAAGTTAGTGGGAAATTAATTTGGGCAAGTCAAGATTATCCCGAGGATTAAG 1575
QY 1493 TGGCCAGCTGTGCAAGCTGTGCGCGGCGCAAGGCGCCTGACCGACATCTGCGCCCTGA 1552

Db 1576 TAAGCAATATTGTAACCTCTTAGAGAAACCAAGCACTAACAGAAGTAATATACCATAA 1635
QY 1553 CCGAGAGCGGAGCTGGAGCTGGCGCGAGAACCGGAGATCTCTGCGGAGCCCGTGCAGG 1612
Db 1636 CAGAAGAAGCAGAGCTAGAACTGGCAGAAACACAGAGATTTCTAAAGAACCAACAGTACATG 1695
QY 1613 GCGTACTACAGACCCAGCAAGGACCTGGTGGCGAGATCAGAAAGCAGAGCCACAGACC 1672
Db 1696 GAGTCTATTATGACCCATCAAGAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 1755
QY 1673 AGTGGACCTACAGATCTACCAGGAGCCCTTCAAGAACTTGAAGACCCGCAAGTACGCCA 1732
Db 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGGAAAAATGCAA 1815
QY 1733 AGATCGCGACCGCCACACCAACAGCAGTGAAGCAGCTGACCCAGGCGCTGCAAGAGATCG 1792
Db 1816 GAATGAGGGGTGCCACACTAATGATGTAACAAATTAACAGAGCAGTGCAAAAAATAA 1875
QY 1793 CCATGGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAGG 1852
Db 1876 CCACAGAAGCATAGTAAATATGGGAAAGACTTCTTAATTTTAAACTGCCATACAAAAGG 1935
QY 1853 AGACCTGGGAGACCTTGTGGACCGACTACTTGGCAGGCCACCTTGGATCCCCGAGTGGAGT 1912
Db 1936 AAACATGGAAACATGTTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGAGT 1995
QY 1913 TCGTGAACACCCCGCTTGTGTAAGCTGTGTACAGCTGTGGAAGAGAGCCCATCATCG 1972
Db 1996 TTGTAAATACCCCTCCCTTAGTGAATATTGTTACCGTTAGAGAAAGAAACCCATAGTAG 2055
QY 1973 GCGCGAGACCTTCTACGTGGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCG 2032
Db 2056 GAGCAGAAACCTTCTATGTAGATGGGCGAGCTAACAGGAGACTAAATTAGGAAAAAGCAG 2115
QY 2033 GCTAGTGACCGACCGGGCGCGCAGAAAGATGTGAGGCTTACCAGGAGACCAACCAACAGA 2092
Db 2116 GATATGTTACTTAATAGAGAAAGACAAAAAGTTGTCAACCTTACTGACACAACAAATCAGA 2175
QY 2093 AGACCGAGCTGCAGGCCATCCAGCTGGCGCTGCAGGACAGCGGAGGAGTGAACATCG 2152
Db 2176 AGACTGATTTACAAGCAATTTATCTAGCTTTGACAGGATTCGGGATTTAGAAGTAAACATAG 2235
QY 2153 TGACCGACGACGAGTACGCTGGGCATCTCCAGCGCCAGCCGACAGAGCAGAGAGG 2212
Db 2236 TAACAGACTCACAATATGCATTTAGAAATCTCAAGACACCAACCATCAAGTGAATCAG 2295
QY 2213 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACCTGAGCTGGG 2272
Db 2296 AGTTAGTCAATCAAAATAATAGACAGTAAATAAAAAAGGAAAGGTCTATCTGGCATGGG 2355
QY 2273 TGCCCGCCCAAGGGATCGGCGCAGCAGCAGCAGATCGAAGCTGGTGAAGCAAGGCA 2332
Db 2356 TACCAGCACAAAGGAATTGGAGGAAATGAACAAAGTAAATAGTCAAGTGTGCTGTGAA 2415
QY 2333 TCCGCAAGGTGCTGTTCTCTGGACGCGCATCGAT 2364
Db 2416 TCAGGAAAGTACTATTTTAGATGAATAGAT 2447

RESULT 6

US-09-735-487-9

; Sequence 9, Application US/09735487

; Patent No. US20020042679A1

; GENERAL INFORMATION:

; APPLICANT: de BETHUNE, Marie-Pierre

; APPLICANT: HERTOGS, Kurt

; APPLICANT: PAUWELS, Rudi

; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE

; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY

; FILE REFERENCE: 1377-125P

; CURRENT APPLICATION NUMBER: US/09/735,487

```

; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)..(489)
; OTHER INFORMATION: gag p6 (52 AA)
; US-09-735-487-9

```

1195	QY	1019	TCACGCTGCCCTGGACGAGGACATTCGCGAAGTACACCGGCCTACACCATTCCTCCAGCACTA
1196	Db	1096	TTTCAGTTCCTCTAGATGAAGACTTTCAGGAAGTATACCTGCATTTACCATACCTTAGTATATA
1197	QY	1079	ACAACGAGACCCCGGCATCCGCTACCCAGTACAACTGCTGCCACAGGCGCTGGAAGGGCA
1198	Db	1156	ACAATGAGACACACAGGGATTAGATATCATGTACATGTGCTTCCACAGGGATGNAAGGAT
1199	QY	1139	GCOCACGATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCCCGCCGCAACC
1200	Db	1216	CACCAGCAATATTCCAAGTAGCATGACAAATACTTTAGAGCCCTTTTAGAAAAACAAATC
1201	QY	1199	CCGAGATCGTGAICTACCA-----GCCCCCTGTGACTGGCGACGGACCTGGAGATCG
1202	Db	1276	CAGACATAGTTATCTATCAATACATGGATGATTGTGATGTAGGATCTGACTTAGAATATG
1203	QY	1253	GCCACGACCGCGCCAAAGATCGAGGAGCTGCGCAAGCACTGCTCGCTGGGGCTTACCA
1204	Db	1336	GCGACATAGACAAAAAATAGAGGAGCTGAGACACATCTGTTGAGTGGGGACTTACCA
1205	QY	1313	CCCCGCAAGAAGCAACCCAGAAGAGCCCCCTTCTCTGTGGATGGGTACGAGCTGCACC
1206	Db	1396	CACCAGACAAAAACATCAGAAAGACCTCCATCTCTTTGGATGGGTATGAATCTCCATC
1207	QY	1373	CCGACAAGTGGACCGTGCAGCCCATCAGCTGCTCCCGAGAAAGGAGAGACTGACCGCTGAACG
1208	Db	1456	CTGATAATGGACAGTAGTACGCTTATGTGCTGCCAGAAAAGACAGCTGGACGTGTCATG
1209	QY	1433	ACATCCAGAGACTGGTGGGCAAGCTGAACCTGGGCCAGCCAGACTCTACCCCGGCATCAAGG
1210	Db	1516	ACATACAGAAGTAGTGGGGNAATTGAAATGGCGCACTCAGATTTACCCAGGATTAAG
1211	QY	1493	TGCGCCAGCTGTGCAAGCTGCTGCGCGCGCCCAAGGCCCTGACCGCAATCGTGCCCTGA
1212	Db	1576	TAAGGCAATTATGTAAACTCTTTAGAGGAAACCAAGACACTAACAGAAAGTAATACCACTAA
1213	QY	1553	CCGAGGAGCGGAGCTGGAGCTGGCCAGAACCCGAGAGATCCTCGCGAGCCCGTGCACG
1214	Db	1636	CAGAAGAAGCAGAGCTAGAACTGGCAGAAAACAGAGAGATTCTAAAAGAACCACTAGTACATG
1215	QY	1613	GCGTGTACTACGACCCCGCAAGGACCTGGTGGCGCGAGATCCAGAAGCAGGGCCACGACC
1216	Db	1696	GAGTGTATTATGCCCATCAAAGACTTAATAGCGAATAACAGACGACGGGCAAGGCC
1217	QY	1733	AGTGGACCTACCGATCTACCGAGAGCCCTTCAAGAACCTGAAAGACCGGCAAGTACGCCA
1218	Db	1756	AATGGACATATCAAAATTTATCAGAGCCATTTAAAAATCTGAAAAACAGGAAAAATATGCAA
1219	QY	1816	GAATGGGGGTGCCACACTAATGATCTAAACAAATTTAACAGAGGCACTGTCAAAAATAA
1220	Db	1793	CCATGTGAGAGCATCGTGCATCTGCGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAAAGG
1221	QY	1876	CCACAGAAAGCATAGTAAATATGGGGAAGACTCCTAAATTTAACTGCCCATACAAAGG
1222	Db	1953	AGACCTGGGAGACCTGTGTGACCGACTACTGGCAGGCCACTGGATCCCCCGAGTGGAGT
1223	QY	1936	AAACATGGGAACATGTTGGACAGAGTATGGCAAGCCACTGGATTCCTGAGTGGAGT

Db 556 AAATGAGTTTCCAGGAAGATGGAACCAAAATGATAGGGGAATTTGGAGGTTTATCA 615
QY 539 AGTGCGCCAGTACGACAGATCTGTGATGAGATCTCGGCGCAAGAAGCCATCGGCACCG 598
Db 616 AAGTAAGACAGTATGATCAGACTACTATAGAAATCTGTGACATATAAGCTATAGGTACAG 675
QY 599 TGCATGATCGGCCCAACCCCGTGAACATCATCGGCCCAACATCTGACCCAGCTGGGCT 658
Db 676 TATTAGTAGNCCTACACCTGTCAACATATTGGAAGAAATCTGTTACTCAGATTGGTT 735
QY 659 GCACCCCTGAATCTCCCATCAGCCCCATCGAGACCGTGCOCGTGAAGCTGAAGCCCGCA 718
Db 736 GCACCTTTAAATTTCCCATATAGCCCTATTGAGACTGTACAGTAAATTAAGCCAGGAA 795
QY 719 TGGAGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGGAGGAGAAATCAAGCCCTGACC 778
Db 796 TGGATGGCCCAAAAGTTAAACAATTTGGCCATTGACAGAGAAATAAAGCATTAGTAG 855
QY 779 CCATCTCGGAGGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGAGAACCCCT 838
Db 856 AAATTTGTACAGATGGAAGAAGGAGGAAATTTCAAAATTTGGGCTTGAATCCAT 915
QY 839 ACAACACCCCGTTCGCCATCAAGAAAGAGGACAGCACCAAGTGGCGCAAGCTGGTG 898
Db 916 ACAATACTCCAGTATTTGCCATAAAGAAAGAACAGACTACTAAATGGAGAAATTTAGTAG 975
QY 899 ACTTCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGTGGGCATCCCC 958
Db 976 ATTTTCAGAGAATTTAATTAAGAGAACTCAAGACTTCTGGGAAAGTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGGCTGAAGAAGAAAGAGCGTGCACCGTGTGACGTGGCGGACCGCTACT 1018
Db 1036 ATCCCGGAGGCTTAAAGAGAAATTAAGTAACTGATGAGTGGGCTGATGCATATT 1095
QY 1019 TCAGCGTGCCTTGAGCAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
Db 1096 TTTTCAGTTCCTTAGATGAAGACTTCAGGAAGTATACCTGCAATTTACCATACTAGTATA 1155
QY 1079 ACAAGGACCCCGGATCCGCTACCAAGTACACAGTGTCTCCCCAGGGCTGGAAGGCA 1138
Db 1156 ACAATGAGACACCGAGGATTTAGATATCAGTACAATGTCTTCCACAGGGATGGAAGGAT 1215
QY 1139 GCGCCAGCATCTCCAGAGCAGCATGACCAAGATCTCGAGCCCTTCCGCGCCCGCAACC 1198
Db 1216 CACCAGCAATATTCAGAGTAGCATGACAAATCTTAGACCTTTTGAARAAACAATC 1275
QY 1199 CCGAGATCGTATACCA-----GGCCCCCTGTAGCTGGGAGCGACCTGGAGATCG 1252
Db 1276 CAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAG 1335
QY 1253 GCCAGCAGCGCCCAAGATCCAGAGCTGCCAAGCAGCTCTGGCTGGGCTTCACCA 1312
Db 1336 GCGAGCATAGAACAAATAGAGGAGCTGAGACACATCTCTTGGAGTGGGAGCTTACCA 1395
QY 1313 CCCCGCAAGAAGCACCAAGGAGCGCCCTTCTGTGGATGGGCTTACGAGCTGCACC 1372
Db 1396 CACCAGACAAAACATCAGAAAGACCTCCATCTTGTGATGGTTTGAACCTCCATC 1455
QY 1373 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCCGGAAGAGAGAGCTGGACCGTGAACG 1432
Db 1456 CTGATAAATGGACAGTACAGCTATAGTGTCTGCCAGAAAAGACAGCTGGACTGTCAATG 1515
QY 1433 ACATCCAGAAGCTGTGGGAGCTGAACTGGCCAGCCAGCATCTACCCCGCATCAGG 1492
Db 1516 ACATACAGAAGTATAGTGGGAAATTTGAATTTGGCAAGTCAATTTACCAGGGATTAAG 1575
QY 1493 TCGCCAGCTGTGAAGCTGTGCGCGCGCCAGCCCTCACCGACATCGTGCOCCTGA 1552
Db 1576 TAAGCAATTATGTAACTCTTAGAGGAACCAAGACACTACAGAAGTAATACCCTAA 1635
QY 1553 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGCGAGATCTCGCGGACCCGTCACG 1612

Db 1636 CAGAAGAGCAGAGCTAGAACTGGCAGAAAAACAGAGAGATTTCTAAAAAGAACCACTACATG 1695
QY 1613 GCGTGTACTAGCACCCAGCAGGAGGACCTGTGGCCGAGATCCAGAAGCAGGCGCACAGACC 1672
Db 1696 GAGTGTATTATGACCCCATCAAAAGACTTAAATAGCAGAAATACAGAGCAGGGGCAAGGCC 1755
QY 1673 AGTGACCTACAGATCTTACAGGAGCCCTTCAAGAACCTTGAAGACCCGCAAGTACGCCA 1732
Db 1756 AATGACATATCAAAATTTATCAGAGGCCATTTAAAAATCTGAAAACAGGAAATATGCNA 1815
QY 1733 AGATGGCAGCGCCACACCAACGACGTGAAGCAGCTGACCGAGCCCTGCAGAAAGATCG 1792
Db 1816 GAATGGGGGTGCCCCACACTAATGATGTAAACAAATTTAACAGAGCGACTGCAAAAAATAA 1875
QY 1793 CCATGGAGAGCATCGTATCTGGGCAAGACCCCAAGCTTCGCGCTGCCATCCAGAAGG 1852
Db 1876 CCACAGAAAGCATAGTATATGGGAAAGACTCTTAATTTAAACTGCCCATACAAAAGG 1935
QY 1853 AGACTGGGAGACCTTGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1912
Db 1936 AAACATGGAAACATGTTGGACAGAGTATTGGCAAGCCACCTGGATTCTTCTGAGTGGAGT 1995
QY 1913 TCGTGAACACCCCCCTCTGTGTAAGCTGTGTACCACTGCTGGAGAGAGGCCCATCATCG 1972
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATTTATGTTACCACTTAGAGAAAGAACCCCATAGTAG 2055
QY 1973 GCGCGGAGACCTTCTACGTGGACGCGCCGCAACCGGAGACCAAGATCGCAGAGCGG 2032
Db 2056 GAGCAGAAACCTTCTATGTAGTGGGCGAGCTAACAGGAGACTAAATTAGGAAAGCAG 2115
QY 2033 GCTAGCTGACCGACCGGGCGGCGGAGAGATCGTGAAGCTGACCCGAGACCCACCAACG 2092
Db 2116 GATATGTTACTAATAGAGAGAACAAAAGTTGTCAACCTTAACGACACACAAATCAGA 2175
QY 2093 AGACGAGCTCGAGGCCATCCAGCTGGCCCTGACGAGCAGCGGAGAGGTGACATCG 2152
Db 2176 AGACTGAGTTACAAGCAATTTATCTAGCTTTGCGAGGATTCGGGATTTAGAAAGTAAACATAG 2235
QY 2153 TGACCGCAGCGAGTACGCTGCGCATCTCCAGCGCCAGCCGACAGAGCAGAGCGG 2212
Db 2236 TACAGACTCAATATGATTTAGGAATCTTCAAGCAGCAACCATCAAGTGAATCAG 2295
QY 2213 AGCTGGTGAACCCAGATCATCGAGCAGCTGATCAAGAAGAGAGGTGTACTGTAGCTGGG 2272
Db 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAAGGAAAGGCTATCTTGGCATGGG 2355
QY 2273 TGCCCGCCCAAGGCGATCGCGGCAAGCAGCAGCATCGACAAGCTGGTGACAGAGGCA 2332
Db 2356 TACCAGCACAAAGGAATTTGAGGAAATGAACAAAGTAGATAAATTAGTCTAGTGTGGAA 2415
QY 2333 TCCGCAAGGTGCTGTCTCTGGACGCGCATCGAT 2364
Db 2416 TCAGAAAGTACTATTTTAGATGGAATAGAT 2447

RESULT 9
US-09-999-183-1
; Sequence 1, Application US/099999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 1
; LENGTH: 4307

; TYPE: DNA									
; ORGANISM: Human immunodeficiency virus type I									
; US-09-999-183-1									
Query Match 44.98; Score 1106.8; DB 10; Length 4307;									
Best Local Similarity 68.2%; Pred. No. 8.9e-176;									
Matches 1618; Conservative 0; Mismatches 732; Indels 22; Gaps 5;									
QY	14	TGGCCGAGGCGCATGAGCAGGCCACCA--CGGCCAACATCTCTGATCGAGCGCAACT	70						
DB	1088	TGGCTGAAGCAATGAGCCAGTACAAATTCAGTACCATTAATGATCGAGAGGCAATT	1147						
QY	71	TCAAGGCGCCCAAGCGCATCATCAAGTGCTTCARCTCGCGCAAGGAGGCCACATCGCC	130						
DB	1148	TTAGGAACCAAGAAAGATGTTAAGTGTTTCAATTTGGCAAGAAAGGGCACACGCCA	1207						
QY	131	GCAACTGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGGCAAGGAGGCCACACAGA	190						
DB	1208	GAATTCAGGGCCCTTAGGAAGAGGGCTGTGGAAATGTGGAAGAGGAAGACACCAAA	1267						
QY	191	TGAAGGACTGCACGAGCGCGCCAGCCCACTTCTTCGCGAGGACCTGGCCCTTCCGCCAGG	250						
DB	1268	TGAAAGATTGTACTGAGACAGAGCTAA-TTTTGTAGGAGATCTGGCCCTTCTACAAG	1326						
QY	251	GCAGGCGCGGAGTTCGCCAGCGAGCAGAACCGCGCCACACAGCCCAACAGCGCGAGC	310						
DB	1327	GGAGGCGCAGGATTTTCTTACAGAGCAGACAGCCAGAGCCACAGCCCAACAGAGAGC	1386						
QY	311	TGAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGGCGCGCGAGCGCCAGGGCA	364						
DB	1387	TTGAGTCTGGGTAGAGACAACACTTCCCCCTCAGAGCAGGAGCGGATAGACAAGGAA	1446						
QY	365	-----CCCTGAACTTCCCCAGATACCCCTGTGGCAGCGCCCTTGGTGAAGCATCAAG	418						
DB	1447	CTGTATCTTTAACTTCCCTCAGTCACTCTTTGGCAACGACCCCTCGTCACATARAAG	1506						
QY	419	TGGCGCCAGATCAAGGAGCCCTGCTGCACACCGCGCGGAGCAGCACCGTGTGGAGG	478						
DB	1507	TAGGGGGCAACTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG	1566						
QY	479	AGATGAGCCTGCCCGGCAAGTGGAGCCCAAGATGTCGCGGCATCGCGCGCTTCATCA	538						
DB	1567	AAATGAGTTTGCCAGGAGATGGAAACCAAAATGATAGGGGAATTGGAGGTTTTATCA	1626						
QY	539	AGGTGCCAGTACGACAGATCCTGATGAGATCTGCGCAAGAGGCCATCGGCACCG	598						
DB	1627	AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGCACATAAAGCTAAGTACAG	1686						
QY	599	TGCTGATCGGCCCCACCCCGTGAACATCATCGCGCCACATCTGACCCAGCTGGGCT	658						
DB	1687	TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTT	1746						
QY	659	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCGGCA	718						
DB	1747	GCACTTAATTTTCCATTAGCCCTATTGAGACTGTACAGTAAATTAAGCCAGGAA	1806						
QY	719	TGACGCGCCCAAGGTGAAGCACTGGCCCTGACCGAGGAGAAAGATCAAGCCCTGACCG	778						
DB	1807	TGATGCGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAGCATTAGTAG	1866						
QY	779	CCATCTGCGAGGAGATGGAGAAGGCGGAGATACCAAGATCGGCCCGGAGAACCCCT	838						
DB	1867	AAATTTGTACAGAGATGGAAAGAGGAGGAAATTTCAAAATTTGGGCCCTGAAATCCAT	1926						
QY	839	ACAACACCCCGTGTTCGCCATCAAGAAAGGACAGACCAAGTGGCGCAAGCTGGTGG	898						
DB	1927	ACAATACTCCAGTATTGCCATAAGAAAGAAACAGACTACTAATGGAGAAATTAGTAG	1986						
QY	899	ACTTCGCGAGCTGAACAGCGCACCCAGGACTTCTGGAGGTGACAGCTGGGCATCCCC	958						
DB	1987	ATTTCAGAAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCA	2046						
QY	959	ACCCCGCGCCTGAAGAGAGAGAGCGGTGACCGTGTGGACGTGGCGGAGCCCTACT	1018						
DB	2047	ATCCCGCAGGGTTTAAAGAAAGAAAAATCAGTAAACAGTACTGATGGGTGATGCATATT	2106						
QY	1019	TCAGCGTGCCTCGAGAGGACTTTCGCAAGTACACCGCTTCACCATCCCCAGCATCA	1078						
DB	2107	TTTCAGTTCCTTAGATGAAGACTTCAGGAAGTATCTGTCATTTACCATACCTAGTATAA	2166						
QY	1079	ACAAGAGAGACCCCGGATCGCTACCAGTACAAGTGTCTCCCCAGGGCTTGGAAAGGCA	1138						
DB	2167	ACAATGAGACACCGAGGATTAGATATCAGTACAATGTCTTCCACAGGGATGGAAAGGAT	2226						
QY	1139	GCCCGACCATCTTCCAGAGCAGCATGACCAAGATCTTCGAGCCCTTCGCGCCCGCAACC	1198						
DB	2227	CACCAGCAATATTCCAAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC	2286						
QY	1199	CCGAGATCTGTATCTACCA-----GGCCCCCTGTAGTGGGAGGACCTCGAGATCG	1252						
DB	2287	CAGCATAGTATCTATCAATACATGATGATTTGTATGTAGGATCTGACTTAGAAATAG	2346						
QY	1253	GCCAGCAGCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGGCTTCACCA	1312						
DB	2347	GGCAGCATAGAACAAAAATAGAGGAGCTGAGAACATCTTTGAGTGGGACTTAGCA	2406						
QY	1313	CCCCGACAAAGAACACCCAGAGGAGCCCTTCTGTGTGATGGGTACGAGCTGCACC	1372						
DB	2407	CACCAGACAAAAACATCAGAAAGAACCTCCATTTCTTGGATGGTTATGAACCTCCATC	2466						
QY	1373	CCGACAGTGGACCGCTGCAGCCCATCGAGCTGCCGCGAGAGGAGAGCTGGACCTGAACG	1432						
DB	2467	CTGATAATGGACAGTAGTAGCCTTAGTGTGCCAGAAAAGACAGCTGGACTGTCAATG	2526						
QY	1433	ACATCCAGAGCTGGTGGGCAAGCTGAACTGGGCGAGCAGAGCTACCCGGCATCAAGG	1492						
DB	2527	ACATACAGAGTGTAGTGGGNAATTTGAATTTGGGCAAGTACAGATTTACCCAGGGATTAAAG	2586						
QY	1493	TGCGCCAGCTGTGCAAGCTGCTGCGCGCGCCAAAGCCCTGACCGACATCGTGCCCTGA	1552						
DB	2587	TAAGGCAATTATGTAAACTCTTAGAGGAAACCAAGCACTAACAGAAAGTAAATACCCTAA	2646						
QY	1553	CCGAGGAGCGGAGCTGGAGCTGGCCGAGACCCGAGAGCTCTCGCGGAGCCCCGTGCAG	1612						
DB	2647	CAGAAGAGCAGAGCTAGAAGCTGGCAGAAAACAGAGAGATCTTAAAGAACACCATAGATG	2706						
QY	1613	GCCTGTACTACGACCCCGAGGACCTGGTGGCGGAGATCCAGAAAGCAGGGCCACCAACC	1672						
DB	2707	GAGTGATTATGACCCATCAAAGACTTAAATACAGAAATACAGAAAGCGGGCAAGGCC	2766						
QY	1673	AGTGGACCTACAGATCTACAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	1732						
DB	2767	AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGAAAAATATCAA	2826						
QY	1733	AGATGGCACCGCCCGCACCAACGAGCTGAGGAGCTGACCGAGGCGCTGCAGAAAGATCG	1792						
DB	2827	GAATGAGGGTGGCCACACTAATGTATTAACAATTTAACAGAGGCGAGTCAAAAAATAA	2886						
QY	1793	CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCCGCTCCCATCAGAAAGG	1852						
DB	2887	CCACAGAACCATAGTAAATATGGGAAGACTCCTAAATTTAAACTGCCCATACAAAGG	2946						
QY	1853	AGACTTGGGAGACCTGGTGACCGACTACTTGGCAGGCCACCTGGATCCCCAGTGGAGT	1912						
DB	2947	AAACATGGGAACATGGTGGACAGAGTATTGGCAAGCCACTGGATTCCTGAGTGGGAGT	3006						
QY	1913	TCGTGAACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGCCCATCATCG	1972						
DB	3007	TTGTTAATACCCCTCCCTTAGTGAATTTATGGTACCAATTTAGAGAAAGAACCCATAGTAG	3066						
QY	1973	GCGCCGAGACCTTCTAGCTGGACGGCGGCCCAACCGCAGAGACCAAGATCGGAAGGCCG	2032						
DB	3067	GAGCAAAACCTTCTTATGTAGATGGGCGACTAACAGGAGACTTAATTTAGGAAAGCAG	3126						
QY	2033	GCTACGTGACCGCGGGCGCGCAAGATCGTGAAGCTGACCGCTGACCGAGACCAACACAGA	2092						


```
QY 1493 TGGCCAGAGCTGTGCAAGCTGTGCGCGCGCAAGGCGGCTGACCGACATCGTGCCCTGA 1552
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3376 TAAGCAATTATGTAAGTCTCTAGAGGAACCAAGACACATAACACTAA 3435
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1553 CCGAGGAGGCGGAGCTGGAGTGGCCGAGAACCCGAGAGATCTCGGAGGCGGCTGAGC 1612
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3436 CAGAGAGCAGAGCTAGAACTGGCAGAAAACAGAGAGTTCTAAAAGACCACTACATG 3495
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1613 GCCTGTACTAGACCCAGCAAGGACCTGTGTGGCGGAGATCCAGAACGAGGCGCACGACC 1672
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3496 GAGTGTATTATGACCCATCAAAAGACTATAGCAGAAATACAGAGGCGGCAAGGCC 3555
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1673 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGGAAGTACGCCA 1732
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3556 AATGGACATATCAAAATTTATCAGAGGCCATTTAAATCTGAAAACAGGAATATGCAA 3615
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1733 AGATGCCAGCGGCCACCAACAGACCTGAAGCAGCTGACGAGGCGGCTGAGAGATCG 1792
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3616 GAATGAGGGGTGGCCACACTAATGATGTAACAATTTAAGAGGCGAGTGCAAAATAA 3675
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1793 CCATGGAGAGCATCGTGTCTGGGGCAAGACCCCAAGTTCCGCGCTGCCATCCAGAGG 1852
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3676 CCACAGAAAGCATAGTAAATGAGGAAGACTCTTAATTTAAACTGCCATACAAAGG 3735
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1853 AGACCTGGGAGACCTGTGTGACCGACTACTGGCAGGCGCACTGTGATCCCGAGTGGAGT 1912
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3736 AAACATGGGAACATGTTGGACAGAGTATTGGCAAGCCACCTGGATTCTCTGAGTGGAGT 3795
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1913 TCGTGAACACCCCGCTCTGGTGAAGCTGTGTACAGCTGGAGAGGAGCCCATATCG 1972
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3796 TTGTTAATACCCCTCCCTCTAGTGAATTTATGTTACCAAGTTAGAGAAAAGCCATAGTAG 3855
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1973 GCGCGGAGACCTTCTAGCTGGAGCGGCGCCCAACCGGAGACCAAGATCGGCAAGGCG 2032
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3856 GAGCAGAACTTCTATGTAGATGGGCAGCTTACAGGGAGACTAAATTAGGAAGACAG 3915
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2033 GCTACGTGACCGACCGGGCGGAGAGATCGTGTGACCTGTGACCGGAGACCAACACAGA 2092
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3916 GATATGTTACTAATAGAGGAAGACAAAAGTTGTCAACCTTAACGTGACACAAATCAGA 3975
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2093 AGACCGAGCTGACGGCCATCCAGCTGCGCTGCGAGGACACCGGAGCGAGGTGACATCG 2152
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3976 AGACTGAGTTACAGCAATTTATCTAGCTTTGAGGAGTTAGAGTAAACATAG 4035
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2153 TGACCCGACACCGAGTACGCTCGGCTATCATCCAGGCGCCAGCGGAGGAGCG 2212
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4036 TAACAGACTCACAATATGCTATAGGATCATTTCAAGCACACACAGATCAAGTGAATCAG 4095
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2213 AGCTGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2272
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4096 AGTTAGTCAATCAATTAATAGACGAGTTAATAAAGGAAAAGGTTCTATCTGCGATGGG 4155
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2273 TGCCCGCCCAAGGCGATCGGCGGCAACGACGAGATCGACAGCTGTTGAGCAAGGGCA 2332
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4156 TACAGCACACAAGGAATTTGAGGAAGAAATGAACAAGTAGATAAATTAGTCAGTGTGGAA 4215
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2333 TCCGCAAGGTGCTGTTCTCGGAGGCGATCGAT 2364
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4216 TCAGGAAAGTACTATTTTAGATGGAATAGAT 4247
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 11

```
US-09-943-286-3
; Sequence 3, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; OTHER INFORMATION: Plasmid.
US-09-943-286-3
```

```
Query Match 42.8% Score 1054.4; DB 10; Length 8933;
Best Local Similarity 55.4%; Pred. No. 4.6e-167;
Matches 1335; Conservative 279; Mismatches 736; Indels 58; Gaps 6;
```

```
QY 14 TGGCGGAGGCATGAGCAGG---CCACCAGGCCAACATCTGATGAGCGCAGCAACT 70
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1200 UGGUGAAGCAAAUGAGCAAGUAAACAGCUAACCAUAAUGAUGCAGAGAGCAAU 1259
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGGGCAAGGAGGCGCACATCGCCC 130
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1260 UUGAGAACCAAGAAAGUUGUUAAGUUUUAAGUUGGCAAGAAAGGGCACACAGCA 1319
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCAAGGAGGCGCCACAGA 190
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 GAAUUGACAGGCGCCUAGGNAAGGCGUUGUUGAAUUGUAGAAAGAGGACACCAA 1379
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 191 TGAAGGACTGACGAGCGCCGCAACTTCTTTCGCGAGGAGCTGCGCTTCCCCCAGG 250
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 UGAAGAUAUUGACGAGACAGGCUAA-UUUUUUAGGGAAGAUUGCGCCUUCUACAAG 1438
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 251 GCAAGGCGCCGAGT-----TCCCCAGCG 274
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1439 GGAAGGCCAGGAAUUUUUUCAGAGCAGACAGGCCAACAGCCCAACUUUUCUAC 1498
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 275 AGCAGAACCCGCGCAACAGCGCCACAGCGCGAGCTGCAGGTGCGGG-----CGACA 328
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1499 AGCAGACAGGCAACAGCCCCCAGCAGAGAGCUUAGGUGGUGUAGAGACAACA 1558
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 329 ACCCCGCGAGGCGCGCGCGAGCGCCAGGCA-----CCCTGAACCTTCCCCCAGA 382
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1559 ACUCGCCUCAGAGCAGGAGCGGUAAGACAAGAACUGUAUCCUUUAACUUCUCUGA 1618
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 383 TCACCTCTGTGGCAGCGCCCTGTGTGAGCATCAAGTGGCGGCGCAGATCAAGAGGCGC 442
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1619 UCACUUCUUGGCAACGACGCCUCUGUCACAAUAAAGUAGGGGCGCAACUAAAGAGCUC 1678
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 443 TGTGGACACCGCGCGCGACACACCGTGTGGAGGAGATGAGCTTCCCGGCAAGTGA 502
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1679 UAUUAGAUACAGGAGCAGAUACAGAUUUAAGAAAUAGAUUUGCCAGGAAGAUUGA 1738
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 503 AGCCCAAGATGATCGGCGGCTCGGCGGCTTCAAGTGGCGGCTGCGGCTGCGGCTGCG 562
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1739 AACCAAAUAGUAGGGGAAUUGGAGUUUUAUAAAGCCAGGAUUGGAAUUAUUUUUCCAUUAGCC 1798
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 563 TGATCGAGATCTCGGCAAGAGGCCCATCGGCACCGTGTGATCGGCCCGCCACCGCTGA 622
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 UCAUAGAAUUCUGUGGACAUAAAGCUAAGUACAGAUUUAUAGUAGGACCUACACCUUCA 1858
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 623 ACATCATCGGCGCGCAACATCTGACCCAGCTGGGCTGACCCCTGAACCTTCCCCATCAGCC 682
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 ACAUAAUUGGAGAAAUUCUGUACUCAGAUUGUUGAUUGGAAUUAUUUUUCCAUUAGCC 1918
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 683 CCATCAGACCGTGCCTGCAAGCTGAAGCCCGCCGATGAGCGGCGCCCGCCCAAGGTGAAGCAGT 742
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CUUUGAGACUGUACCGAAUUAUAAAGCCAGGAUUGGAAUUAUUUUUCCAUUAGCC 1978
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 743 GGCCCTGACCGAGGAGAGATCAAGGCGCTGACCGCATCTGCGAGGAGATGAGGAGG 802
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 GGCAUUGACAGAGAGAAAAUAAAGCAUUAUAGUAGAAUUUUGUACAGAAUUGGAAAGG 2038
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 803 AGGGCAAGATCACCAAGATCGGCGCGCGAGAAACCCCTTACAACACCCCGCTGTTCGCCATCA 862
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 1320 GAAUUGCAGGGCCCCUAGAAAAAGGGGUGUUGGAAUUGUGAAAGGAGACACAAA 1379
QY 191 TGAAGACTGCACGAGCGCCAGGCAACTTCTCCGCGAGGACCTGGCTTCCCCCAGG 250
Db 1380 UGAAAGUUGUACUGAGACAGCGCUAA-UUUUUUAGGAAGAUUGGCUUCCUACAAG 1438
QY 251 GCAAGCCCGCGAGT-----TCCCCAGG 274
Db 1439 GGAAGCCAGGGAUUUUCUACAGACGACACAGACCCACCCACCAUUUCUACAG 1498
QY 275 AGCAGAACCGCGCCAAACAGCCACACAGCGCGAGTGCAGTGCAGTGCAGG 328
Db 1499 AGCAGACGAGACGACAGCGCCACACAGAGAGAGAGUUCAGGUGGUGAGACAACA 1558
QY 329 ACCCCGCGACGAGCGCGCGCGCGCGCGCGCA-----CCCTGAAGTCCCGCCAGA 382
Db 1559 ACUCCCCUACAGACGAGGCGGAGUAGACAGGAACUGUUCUUUAACUCCUACAGA 1618
QY 383 TCACCCCTGTGGACGCCCCCTTGTGTGAGCATCAAGGTGGCGCGCCAGATCAAGGAGGCC 442
Db 1619 UCACUUCUUGGCAACGACCCUGUCACAAUAAAGAUAGGGGGCAACUAAAGGAAGCUC 1678
QY 443 TGTGTACACCGCGCGCGAGCACCGTGTGTGAGAGAGATGACCTTGGCGGCAAGTGA 502
Db 1679 UAUUAGAUACAGGAGCAGAGUACAGUUAUAGAAAUAGAUUUGCCAGGAAGUUGGA 1738
QY 503 AGCCCAAGATGATCGCGGCATCGCGGCTTTCATCAAGGTGCGCCAGTAGACACAGATCC 562
Db 1739 AACCAAAUAGUAGGGGAUUUGGAGGUUUUAUCAAAGUAGACAGAUAGUACAGAUAC 1798
QY 563 TGATCGAGATCTCGGCAAGAGGCCATCGGCACCGTGTGTGTCGCCCGCCACCGCGTGA 622
Db 1799 UCAUAGAAUUCUGGACAAUAAAGUAGUACAGUUAUAGUAGGACCUACACCUUGUA 1858
QY 623 ACATCATCGGCCCGACATGCTGACCCAGCTGGCTGGCTGCACCTGGAATCTCCCGATCAGCC 682
Db 1859 ACAUAAUUGGAAGAAUUCUGUUGACUACAGUUGGUGUCCUUAUUAUUCUAGGCC 1918
QY 683 CCATCGAGACCGTGCCTGTAAGCTGAAGCGCGGATGGAGCGCCCAAGGTGAAGCAGT 742
Db 1919 CUUUGAGACUGUACCAUAAUUAAGAGGAAUUAAGCCAGGAUUGGCCCAAAAGUUAACA 1978
QY 743 GGCCCTGACCGAGGAGAGATCAAGCCCTGACGCCCATCTGCGAGGAGATGGAGAAG 802
Db 1979 GGCCAUUGACAGAAAGAAUAAAGCAUUAUAGAAUUAUUGACAGAAUUGGAAAG 2038
QY 803 AGGCAAGATCACCAGATCGGCCCGCGAGACCCCTACACACCCCGCTGCGCCATCA 862
Db 2039 AAGGAAAUUUAUAAAUUUGGCCUUGAGAAUCCAUACAACUCCAGUAUUGCCAUAA 2098
QY 863 AGAAGAGGACAGACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCA 922
Db 2099 AGAAAAGACAGUACUAAUAGGAGAAAUUAGUAGUUAUUCAGAACUUAUAGAGAA 2158
QY 923 CCAGAGATCTGGGAGGTGACAGTGGGATCCCGCCACCGCGCGCTGGAAGAGAA 982
Db 2159 CUCAAGACUUCUGGGAUUGUUAUAGGAUUAACCAUCCCGCGGUGUUAUAAAGAAA 2218
QY 983 AGAGCTGACCGTGTGGAGTGGCGAGCGCCTACTTCAGCTGCGCTTGGAGGAGACT 1042
Db 2219 AUCAGUAACAGUACUGGAGUGGUGUAGUAGUUAUUAUUAUUAUUAUUAUUAUUA 2278
QY 1043 TCCGCAAGTACACCGCTTCCACCATCCAGCATCAACACAGAGACCGCGCGATCCGCT 1102
Db 2279 UCAGGAAGUUAUCUGAUUACCAUACUUAUUAUUAUUAUUAUUAUUAUUAUUA 2338
QY 1103 ACCAGTACAACTGTGTGCCCCAGGGTGTGAAGGCGACCGCCAGCATCTTCCAGAGCA 1162
Db 2339 AUCAGUACAUGUGUUCUCCACAGGGAUGGAAAGGAUACACAGCAUUAUUCCAAAGUAGCA 2398
QY 1163 TGACCAAGATCTGTGGAGCGCTTCCGCGCGCGCAACCCCGAGATCTGTACTACCA----- 1217
Db 2399 UGACAAAAUUCUAGAGCCUUUUAUAAACAAUUAUUAUUAUUAUUAUUAUUAUUA 2458

QY 1218 -GGCCCCCTGTACGTGGCAGCGACCTTGAGATCGGCAGCACCGCCCAAGATCGAGG 1276
Db 2459 UGGAUUAUUGUAGUAGAUUGACUUAAGAAUAGGCGAGCAUAGAAACAAAUAGAGG 2518
QY 1277 AGTGGCGAAGCACCTGCTGCGCTGGGGTTCACCCCGCCGACAAAGACACCAAGG 1336
Db 2519 AGCUGAGACAAACUUCUUGAGGUGGGGACUUAACACACAGACAAAAACAUCAGAAAG 2578
QY 1337 AGCCCCCTTCTGTGTGGATGGGTACGAGTGCACCCCGACAAAGTGGACGCTGCAGCCCA 1396
Db 2579 AACCCUUAUUCUUGGAGUGGUUAUGAACUCCUACUGAAUAAUAGGACAGUACAGCCUA 2638
QY 1397 TCGAGTGCCTCCGAGAGAGAGCTGACCGTGAACACATCCCAAGAGCTGGTGGCAAGC 1456
Db 2639 UAGUGUGCCAGAAAAGACAGCUGGACUGUACAUAGACAUAGAUUAGUGGGNAU 2698
QY 1457 TGAACCTGGCGACCGACATGTTACCCCGGCATCAAGGTGCGCCAGCTGTGCAAGCTGCTGC 1516
Db 2699 UGAAUUGGCAAGUCACAGAUUUACCCAGGGAUUAAGUAAGGCAAUUAUUAUAAUCUUA 2758
QY 1517 GCGGCGCCAGGCGCTGACCGACATGTCGCCCTGACCGAGGAGCGCGAGCTGGAGCTGG 1576
Db 2759 GAGGAACCAAGACACUACAGAAUAUACCAUACAGAAAGAGCAGACUAGACUGG 2818
QY 1577 CCAGAACCGCGAGATCTCTGCGCGAGCCCGTGCACGCGCTGTACTACGACCCCGCAAGG 1636
Db 2819 CAGAAAACAGAGAGAUUUAUAAAGAACACAGUACAUAGGAGUUAUUAUAGCCCAUAAAG 2878
QY 1637 ACCTGTGGCGGAGATCCAGAACGAGGCGCACGACAGTGGAGCTTCCAGATCTTACCAGG 1696
Db 2879 ACUUAUAGCAGAAUACAGAGAGCGGCGAGGCCAAUAGGACAUUCAAUUAUUAUUA 2938
QY 1697 AGCCCTTCAAGAACCTGAAGACCGCGCAAGTACGCCAGATGCCGCGCCCGCACCAACG 1756
Db 2939 AGCAUUAUAAAUUAGAAACAGGAAAUUAUGCAAGAUAGGUGGUGGCCACACUAAUG 2998
QY 1757 ACGTGAAGCAGCTGACCGCGCGCGTGCAGAAAGATCGCATGTGAGAGCATCTGTCTGGG 1816
Db 2999 AUGUAAAACAAUUAACAGAGCGAGUGCAAAAUAUAAACACAGAAAGCAUAGUAAUAGG 3058
QY 1817 GCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAGGAGACTGGGAGACCTGTGTGACCG 1876
Db 3059 GAAAGACUCCUAAAUUUAUAAUACUCCCAUAAAGAAACAUUGGAAACAUUGGUGGACAG 3118
QY 1877 ACTACTGGGAGCGCACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCGCCCTGTGA 1936
Db 3119 AGUUAUGGCAAGCCACCGUAGUUCUGAGUGGAGUUUUAUUAUUAUUAUUAUUAUGA 3178
QY 1937 AGCTGTGGTACCAAGCTGGGAAGGAGCCCATCATCGCGCGCGAGACCTTCTAGCTGGACG 1996
Db 3179 AUUAUUGUACCAUUAAGAAAGAACCCAUAGUAGGAGCAGAAACCUUAUUAUUAUGA 3238
QY 1997 GCGCCGCAACCGCGAGACCAAGATCGGCAAGCGCGCTACGTGACCGAGCGGCGCGCG 2056
Db 3239 GGCAGCUAACAGGAGACUAAAUUAGGAAAGCAGGAUUAUUAUUAUUAUUAUUAUUA 3298
QY 2057 AGAAGATCTGACCTGACCGGACACCAACAGAAAGCAGAGCTGCGAGGCTTCCAGC 2116
Db 3299 AAAAGGUGUCCCGCUUAUACACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3358
QY 2117 TGCCCTTGCAGGACAGCGGAGGTGAACATCTGTGACCGCAGCAGCTAGCGCTGG 2176
Db 3359 UACUUAUGCAGGAUUAUAGGAUUAAGAAUAAUUAUUAUUAUUAUUAUUAUUAUUA 3418
QY 2177 GCATATCCAGGCGCGCGAGCAGAGCAGCAGCTGCTGTGTAACCAATCATCAGC 2236
Db 3419 GAAUUAUUAAGCAGACACCAAGAAUUAAGUUAUUAUUAUUAUUAUUAUUAUUAUUA 3478
QY 2237 AGCTGATCAAGAGGAGAGGTGTACTGAGTGGGTGCGCGCCCGCACAGGAGCTCGCG 2296
Db 3479 AGUUAUAAUAAAGGAAAGGUCUUCUGGCAUGGUGUACCAAGCAGACAAAGAAUUGGAG 3538


```

Db 2819 CAGAAAAACAGAGAGUUCUAAAAAGAACACAGUACAUAGGAGUUAUUAUAGCCCAUCAAAG 2878
QY 1637 ACCTGTTGGCCGAGATCCAGAACGCGCCAGCACAGTGGACCTACCAGATCTACCAGG 1696
Db 2879 ACUUAUAGCAGAAUACAGAACGCGGAGCCCAUAGGACAUAAUUAUUAUUAUUAUUAUUA 2938
QY 1697 AGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCGCCACACCAACG 1756
Db 2939 AGCCAUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2998
QY 1757 ACCTGAAGACGCTACCGAGGCGCTGCAGAAATCGCCATGGAGAGCATCGTATCTGGG 1816
Db 2999 AUGUAAACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3058
QY 1817 GCAAGACCCCAAGTTCGCGCTCCCATCCAGAAAGAGAGACCTGGAGACCTGTGGACCG 1876
Db 3059 GAAGACUCCUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3118
QY 1877 ACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCGCCCTGGTGA 1936
Db 3119 AGUUAUGGCAAGCACCGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3178
QY 1937 AGCTGTGTACGAGCTGGAGAAAGAGCCCATCATCGGCGCCGAGACCTTCTACGTGAGC 1996
Db 3179 AUUAUUGGUUACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3238
QY 1997 GCGCCGCAACCGCAGACCAAGATCGGCAAGCGCGCTACGTGACCGACCGGCGCGG 2056
Db 3239 GGCACGUAACAGGAGACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3298
QY 2057 AGAAGATCGTGAAGCTGACCGAGACCAACACAGAACGCGAGCTGCGAGGCTCCAGC 2116
Db 3299 AAAAGGUGUCCGCCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3358
QY 2117 TGCCCTGAGGACGCGGAGCGAGGTGAAATCTGTGACCGACGCGCAGTACGCCCTGG 2176
Db 3359 UAGCUUUGCAGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3418
QY 2177 GCATCATCAGGCGCCGACCAAGAGGAGAGCGAGCTGTGTGAACACAGATCATCGAGC 2236
Db 3419 GAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3478
QY 2237 AGCTGATCAAGAGAGAGAGTGTACTGAGTGGTGGCGCCGACCAAGGCGATCGGCG 2296
Db 3479 AGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3538
QY 2297 GCAACGAGCAGATCGACAGCTGCTGAGCAAGGCGATCCGCAAGGTGCTTCTTGGACG 2356
Db 3539 GAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3598
QY 2357 GCATCGAT 2364
Db 3599 GAUAUUAU 3606
```

RESULT 14

```

US-09-798-675-4
; Sequence 4, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798,675
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186,364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251,083
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
```

```

; LENGTH: 9544
; TYPE: DNA
; ORGNISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct of vaccine vector pGA2 and insert JS2 expressing cla
; OTHER INFORMATION: HIV-1 VL
; NAME/KEY: misc.feature
; LOCATION: (106)..(1641)
; OTHER INFORMATION: encoding inner core protein
; NAME/KEY: misc.feature
; LOCATION: (1401)..(3617)
; OTHER INFORMATION: encoding viral rproteins including reverse transcriptase
; NAME/KEY: misc.feature
; LOCATION: (3708)..(6334)
; OTHER INFORMATION: induces high level of virus particles
; NAME/KEY: misc.feature
; LOCATION: (3847)..(6518)
; OTHER INFORMATION: encodes Rev regulating transfer of RNA to cytoplasm
; NAME/KEY: misc.feature
; LOCATION: (4102)..(6660)
; OTHER INFORMATION: encoding ADA Env
; NAME/KEY: misc.feature
; LOCATION: (6672)..(9544)
; OTHER INFORMATION: Vaccine vector pGA2
; US-09-798-675-4

Query Match 42.5%; Score 1048; DB 10; Length 9544;
Best Local Similarity 66.9%; Pred. No. 5 3e-166;
Matches 1610; Conservative 0; Mismatches 740; Indels 58; Gaps 6;

QY 14 TGCCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACATCTCTGATGCGCAGCAACT 70
Db 1193 TGCTGAAGCAATGAGCAAGTAACAAATACAGTACCAATATGATCAGAGAGGCAATT 1252
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTTCACCTGCGCGCAAGGAGGCGCACATCGCCC 130
Db 1253 TTAGGAACCAAGAAAGAGTGTAAAGAGCTTCAATAGCGCAAGAGGCGCACAGCCA 1312
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCGCACAGA 190
Db 1313 GAATTCGAGGCGCCCTAGGAAAGAGGCGAGCTGGAAAGCGGNAAGGAGGACACCAA 1372
QY 191 TGAAGACTGTCACCGAGCGCCGCAACTTCTTCCGCGAGGACCTGGCCTTCCCCCAGG 250
Db 1373 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTATTAGGAAGATCTGGCCTTCTCTACA 1431
QY 251 GCAAGGCGCGCGAGT-----TCCCCAGCG 274
Db 1432 GGAAGCGCCAGGGAATTTTCTCAGAGCAGACACAGCCACAGCCCATTTCTTCAG 1491
QY 275 AGCAGAACCCGCGCCCAACAGCCCCCAGCGCGAGCTGCGAGTGGCGCGG-----CGACA 328
Db 1492 AGCAGACAGAGCCACAGCCCCCAGCAGAGAGAGCTTTCAGTCTGGGTAGAGACAACA 1551
QY 329 ACCCCGCGAGGAGCGCGCGCGCGAGCCAGGCGCA-----CCCTGAACTTCCCCCAGA 382
Db 1552 ACTCCCCCTCAGAAGCAGGAGCGAGTAGACAAGAACTGTATCTTTAACTTCCCTCAGA 1611
QY 383 TCACCTCTGTGAGCGCCCCCTTGTGAGCATCAAGTGGCGCGCCAGATCAAGGAGGCC 442
Db 1612 TCACCTTTGGCAACGACCCCTCTGCATAAAGATAGGGGGGCACTAAAGAGAGCTC 1671
QY 443 TGCTGGACACCGCGCGCGAGCACCGTGTGGAGGAGATGAGCCTGCGCGGCAAGTGA 502
Db 1672 TATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATGAGTTGCCAGGAGATGA 1731
QY 503 AGCCCAAGATGATGCGCGGCGATCGCGGCTTCATCAAGTGCGCCAGTACGACAGATCC 562
Db 1732 AACCAAAAATGATAGGGGAAATTGGAGGTTTATCAAAAGTAAGACAGTATGATCAGATAC 1791
QY 563 TGATCGAGATCTCGGCAAGAGGCCATCGGCACCGTGTGATCGGCGCCACCCCGCTGA 622
Db 1792 TCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTCTCA 1851
```

QY 623 ACATCGCTGGCGCCCAACATGCTGACCCAGCTGGGTGACCCCTGAACCTGCCCATCAGCC 682
DB 1852 ACATAATTTGAAGAAATCTGTTGACTCAGATTGGTTGCACATTAAATTTCCCATAGCC 1911
QY 683 CCATCGACACGCTGCCGTGAAGCTGAAGCCGGCATGGAGCGCCCAAGGTGAAGCAGT 742
DB 1912 CTATTGAGACTGTACCACTAAATTTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAAT 1971
QY 743 GGCCCTCACCAGGAGGAGAACTCAAGGCCCTGACCCGCATCGCCGATCGCAGAGATGGAGAGG 802
DB 1972 GGCCATTGACAGAGAAAAATAAAGCATTTAGTAGAAATTTGTACAGAAATGGAAAAAG 2031
QY 803 AGGCAAGATCACCAGATCGGCCCGGAGAACCCCTACAAACACCCCGCTGTTCGCCATCA 862
DB 2032 AAGGAAAAATTTCAAAATTTGGGCTGAGATCCATACAAATATCCAGTATTTGCCATAA 2091
QY 863 AGAAGAGGACAGCACCAAGTTGGGCAAGCTGGTGAGCTTCGGGAGCTGAACAAAGCCCA 922
DB 2092 AGAAAAAGACAGTACTAAATGGAGAAATTTAGTAGATTTTCAGAGAACTTAATAGAGAA 2151
QY 923 CCCAGGACTCTGGGAGGTGACGTGGGCATCCCCACCCCGCGGCTTGAAGAGAGA 982
DB 2152 CTCRAGACTTCTGGAGTTCAATTTAGGAATPACCACATCCCGCGAGGGTTAAAAAGAAAA 2211
QY 983 AGAGCGTGACCGTCTGACGCTGGGCGACGCTACTTTCAGCGTGGCCCTGGACGAGGACT 1042
DB 2212 AATCAGTAACAGTACTGATGTGGTGATGCATATTTTTCAGTTCCCTTAGTAGAAGACT 2271
QY 1043 TCCGCAAGTACACCGCTTACCATCCCGACGATCAACAAGAGACACCCCGGCTATCCGCT 1102
DB 2272 TCAGGAATTAATGCAATTTACCATACCTAGTATAAACAATGAGACACACGAGGATTAGAT 2331
QY 1103 ACCAGTACAACGCTGTGCCCGACGCTGGAGGCGAGCCCGACGATCTTCAGAGCAGCA 1162
DB 2332 ATCAGTACAATGCTTCCACAGGATGGAAGGATCACCAAGCAATATTCCAAAGTAGCA 2391
QY 1163 TGACAAAGATCCTGGAGCCCTTCGCGCCCGCAACCCGAGATCGTGATCTACCA - - - - 1217
DB 2392 TGACAAAAATCTTAGAGCTTTTAAAAAACAATCCAGACATAGTTATCTCAATACA 2451
QY 1218 -GGCCCCCTGTAGTGGGACGACCTGGAGATCGGCGACGACCGCCCAAGATCGAGG 1276
DB 2452 TGAACGATTTGTATGTAGGATCTGACTTAGAATAATAGGCGCATAGAACAAAAATAGAGG 2511
QY 1277 AGCTGCGCAAGCCTGTGCTGGCTGGGCTTCACACCCCGCAAGAACACAGAGG 1336
DB 2512 AGCTGAGACACATCTGTTGAGTTGGGACTTACCACCCAGACAAAAAATCAGAAAG 2571
QY 1337 AGCCCCCTTCTGTGATGGGCTACGAGCTGCACCCCGCAAGTGGACCTGCGAGCCCA 1396
DB 2572 AACCTCCATTCCTTGGATGGGTTATGAATCCATCTGTATTAATGGACATCAGACCTA 2631
QY 1397 TCGAGCTGCCGAGAGGAGAGCTGGACCGTGAACGATCCCAAGCTGGTGGGCAAGC 1456
DB 2632 TAGTGCTGCCAGAAAAAGACAGCTGGACTGTCAATGATACATACAGAAAGTTAGTGGGAAAT 2691
QY 1457 TGAATGGGCGAGCCAGATCTACCCCGCATCAAGGTGCGGCGAGCTGTGCAAGCTGCTGC 1516
DB 2692 TGAATCCGCAAGTCAGATTACCAGGGATTAAAGTAAGCAATTAATGTAACCTCCTTA 2751
QY 1517 GCGGCGCAAGGCGCTGACCGACATCTGCGCCCTGACCCGAGGCGCGAGCTGGAGCTGG 1576
DB 2752 GAGGAACCAAGCACTAACAGAACTAATACCATAACAGAGAGAGAGAGTGAAGCTGG 2811
QY 1577 CCGAGAACCGGAGATCTTGGCGAGCGCTGACGCGGTGTACTAGACCCCGCAGAGG 1636
DB 2812 CAGAAAAACAGAGATTTCTAAAAAGAACCGAGTACATGAGGTGTATTATGACCCCAAAAAG 2871
QY 1637 ACCTGGTGGCGGAGATCCAGAGCAGGCGCCAGCAGCTGAGCTACAGATCTACAGG 1696
DB 2872 ACTTAATAGCAAAATACAGAGCAGGCGGCAAGGCCAATGACATATCAAAATTTATCAAG 2931

QY 1697 AGCCCTTCAAGAAACCTGAAGACCGCGCAAGTACGCCAAGATCGCACCGCCACACCAAGC 1756
DB 2932 AGCCATTTAAATCTGAAAAACAGAAATATGCAAGATGAGGGTGGCCCAACATAATG 2991
QY 1757 AGCTGAAGCAGCTGACCGAGGCGGTGCAAGATGCCATGGAGAGCATCGTGATCTGGG 1816
DB 2992 ATGTAAAAACAATTAACAGAGGCGAGTCAAAAAATAACCAACAGAAAGCATAGTAATATGG 3051
QY 1817 GCAAGACCCCAAGTTCGCCCTGCCATCCAGAACGAGACCTGGGAGACCTGGTGAGACCG 1876
DB 3052 GAAAGACTCTTAAATTTAACTTACCCATACAAAAGGAAACATGGGAAACATGGTGACAG 3111
QY 1877 ACTACTGCGAGCCCACTGGATCCCGGAGTGGAGTTCTGTGAACACCCCGCCCTGGTCA 1936
DB 3112 AGTATTGCAAGCCACTGGATTCTCTGAGTGGAGTTGTTAATACCCCTCTCTTAGTGA 3171
QY 1937 AGCTGTGTACAGCTGGAGAGGAGCCCATCATCGGCCCGGAGACCTTCTACGTGGAGC 1996
DB 3172 AATTATGTTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAGATG 3231
QY 1997 GCGCCCAACCCCGAGACCAAGATCGGCAAGCGCGGCTACGTGACCGAGCGGGCGCGC 2056
DB 3232 GGGCAGCTAACAGGGAGACTAAATTTAGGAAAAAGCAGGATATGTTACTAACAAAGGAAGAC 3291
QY 2057 AGAAGATCGTGAGCCTGACCGAGAGACCAACCAAGAACCGAGCTGCAGGCGCATCCAGC 2116
DB 3292 AAAAGTTGTCCCTTAATACTACACAAACATCAGAAAACTCAGTTACAAACATTTATC 3351
QY 2117 TGGCCCTCGAGAGACGCGCAGCGAGGTGAACATCGTGACCGCAGCGAGCTACGCGCTGG 2176
DB 3352 TAGCTTTGCGAGATTCAGGATTTAGAAAGTAAACATAGTAACAGACTCACAATATGCATTAG 3411
QY 2177 GCATCATCCAGGCCCGCAGCCGACAGAGCGGAGAGCGAGCTGTGTGAACAGATCATCGAGC 2236
DB 3412 GAATCATTCACAGCACACCAAGTAAAGTGAATCAGATTAAGTCAATCAAAATTAATAGAGC 3471
QY 2237 AGCTGATCAAGAAAGAGAGGTGTACCTGAGCTGGTGGTGGCCCGCCACAGGCGCATCGCG 2296
DB 3472 AGTTAATAAAGAAAGGAAAGGTCTATCTGCGATGGGTACCGACACAAAGAAATTTGGAG 3531
QY 2297 GCAAGCAGCATGCAAGCTGTGTGCAAGGGCATCGCAAGGTGCTGTCTCTCTGGAGC 2356
DB 3532 GAAATGAACAAGTAGATAAATTAGTCAGTGTCTGCTGGAATCAGGAAAAATACTATTTTTAGATG 3591
QY 2357 GCATCGAT 2364
DB 3592 GAATAGAT 3599
RESULT 15
US-09-798-675-5
; Sequence 5, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798,675
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186,364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251,083
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 9918
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct of vaccine vector pGAL and vaccine insert expressing
; OTHER INFORMATION: ade B HIV-1 Gag-Po
; NAME/KEY: misc_feature

; LOCATION: (106)..(1641)
; OTHER INFORMATION: encodes proteins of viral inner core
; NAME/KEY: misc.feature
; LOCATION: (1401)..(3617)
; OTHER INFORMATION: encodes viral proteins but not integrase
; NAME/KEY: misc.feature
; LOCATION: (3708)..(5715)
; OTHER INFORMATION: regulates high-level production of HIV genes
; NAME/KEY: misc.feature
; LOCATION: (3847)..(5944)
; OTHER INFORMATION: encodes protein regulating the transfer of RNA to cytoplasm
; NAME/KEY: misc.feature
; LOCATION: (3939)..(4181)
; OTHER INFORMATION: gene participates in viral assembly and budding
; NAME/KEY: misc.feature
; LOCATION: (4099)..(4941)
; OTHER INFORMATION: encodes truncated form of viral coat protein
; NAME/KEY: misc.feature
; LOCATION: (6098)..(9918)
; OTHER INFORMATION: vaccine vector pGAL
US-09-798-675-5

Query Match 42.5%; Score 1048; DB 10; Length 9918;
Best Local Similarity 66.9%; Pred. No. 5.3e-166;
Matches 1610; Conservative 0; Mismatches 740; Indels 58; Gaps 6;

QY 14 TGGCCGAGCCATGAGCCAGG---CCACCAGCCCAACATCTCTGATCGAGCGCAGCAACT 70
DB 1193 TGGCTGAAGCAATGAGCCAAAGTAAACAAATACAGCTACCATAATGATGCGAGAGGCAATT 1252

QY 71 TCAAGGGCCCCCAAGCGCATCATCAAGTGCCTTCAACTGCGGCAAGGAGGCGCACATGCCCC 130
DB 1253 TTAGGAACCAAGAAAGATGGTTAAGAGCTTCAATAGCGGCAAGAGGCGCACACAGCCA 1312

QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGGTGTGCTGGAAGTGGCGGAAGGAGGCGCCACAGA 190
DB 1313 GAAATTGCGAGGGCCCTTAGGAAAGAAAGGCGAGCTGGAAAGGAAAGGAAAGGACACCAA 1372

QY 191 TGAAGGACTCACCGAGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCGCTTCCCGCAGG 250
DB 1373 TGAAGATTGCTGAGAGACAGGCTAA-TTTTITAGGGAAGATCTGGCCCTTCTCTACAG 1431

QY 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
DB 1432 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGCCACAGCCGCCACCATTTCTTCAG 1491

QY 275 AGCAGAACCCGCGCAACAGCCCGCCAGCGCGAGCTGCGAGTGGCGGG-----CGACA 328
DB 1492 AGCAGACCGAGGCCAACAGCCCGCCAGAGAGAGCTTTCAGGTCTGGGGGTAGAGACAACA 1551

QY 329 ACCCCGCGAGCGGCGCGCGCGAGCGCCAGGCA-----CCCTGAAGCTTCCCGCAGA 382
DB 1552 ACTCCCCCTCAGAGCAGGAGGAGCGATAGACAGGAAGTGTATCTTTAACTTCCCTCAGA 1611

QY 383 TCACCTCTGTGGCAGCGCCCTGCTGAGCATCAAGTGGCGCGCCAGATCAAGGAGGCC 442
DB 1612 TCACTCTTTGGCAACGACCCCTCGTCACATAAAGATAGGGGCGCACTAAAGAGAGCTC 1671

QY 443 TGCTGGACACCGCGCGCGAGCACCGTGTGGAGGAGATGAGCGCTGCGCGCAAGTGA 502
DB 1672 TATTAGATACAGGAGCAGATGATACAGTATTAGAAGAAATCAGTTTGGCCAGGAAGATGA 1731

QY 503 AGCCCAAGATGATCGCGGATCGCGGCTTCATCAAGTGGCGCCAGTACGACAGATCC 562
DB 1732 AACCAAAATGATAGGGGAATTTGAGGTTTATCAAGTAAGACAGATGATGATCAGATAC 1791

QY 563 TGATCTGAGATCTCGGCAAGAGGCCATCGCGACCGCTGCTGATGGCGCCCGCCCGCGTGA 622
DB 1792 TCATAGAAATCTGTGGACATAAAGCTATAGGTACAGTATTAGTAGGACCTACACCTGTCA 1851

QY 623 ACATCATCGCGCGAGCATGCTGACCCAGCTGGCGTGCACCCCTGGAAGTTCGCCATCAGCC 682
DB 1852 ACATAATTGGAAGAAATCTGTTGACTCAGATGTTGCTGACTTTTAAATTTTCCCATTAGCC 1911

QY 683 CCATCGAGACCGTGCCTGAACTGAAAGCCCGCATGGACCGGCCCAAGGTGAAGCAGT 742
DB 1912 CTATTGAGACTGTACCAGTAAATTAAGCCAGGAATGGATGGCCCAAAATTTAAACAAT 1971

QY 743 GGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCGCCATCTCGGAGGAGATGAGAGAG 802
DB 1972 GGCCATTGACAGAAGAAAAATAAAGCATTAGTAGAAATTTGTACAGAAATGAAAGG 2031

QY 803 AGGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACACACCCCGCTTCGCCATCA 862
DB 2032 AAGGGAATAATTTCAAAATTTGGCCCTGAGAAATCCATACAATCTCCAGTATTTGCCATA 2091

QY 863 AGAAGAGGACGACACCAAGTGGCGCAAGCTGGTGGACTTCCCGGAGCTGAACAAGCGCA 922
DB 2092 AGAAAAAGACAGTACTAATATGGAGAAAAATTAGTAGATTTTCAGAGAACTTTAATAAGAGAA 2151

QY 923 CCAGGACTTCTGGGAGGTGCACTGGGCATCCCGCCACCCCGCGGCTGAAGAAGAGA 982
DB 2152 CTCAGGACTTCTGGGAAGTTCAATTAGGAATACCAATCCCGAGGGTTAAAAAAGAAA 2211

QY 983 AGAGCGTGACCGTGTGGAGCTGGCGACCGCTTACTTCAGCGTCCCGCTGGACGAGCT 1042
DB 2212 AATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTCAGTTCCCTTAGATGAAGACT 2271

QY 1043 TCCGCAAGTACACCGCTTCCACCATCCCGAGCATCAACAACAGAGACCCCGGATCCGCT 1102
DB 2272 TCAGGAAATGATGCACTTACCATACCTAGTATAAACATAGACACACCGGATTAGAT 2331

QY 1103 ACCAGTACAAAGTGTCTGCCCGGAGGTGGAAGGCGACGCCCGAGCATCTTCAGAGCAGCA 1162
DB 2332 ATCAGTACAATGTCTTCCACAGGGATGGAAGGATCACCAGCAATATTTCAAAGTAGCA 2391

QY 1163 TGACCAAGATCTTGGAGCGCTTCCGCGCCGCAACCCCGAGATCGTGTATACCA----- 1217
DB 2392 TGACAAAAATCTTAGAGCCCTTTAAAAAACAATAATCCAGACATAGTTATCTCATATACA 2451

QY 1218 -GGCCCCCTTCTAGTGGCGAGGACCTGGAGATTCGCCAGCACCGCCCAAGATTCGAGG 1276
DB 2452 TGAACGATTTGTATGATAGGATCTGACTTTAGAAATAGGCGAGCATAGAACAAAAATAGAGG 2511

QY 1277 AGTGGCAGAGCAGCTGCTGCGGTGGGGCTTACACCCCGCGAGCAAGAGCAGAGAG 1336
DB 2512 AGCTGAGACAACTCTGTTGAGGTGGGGACTTACCACACAGACAAAAAATCAGAGAG 2571

QY 1337 AGCCCCCTTCTGATGGGTGAGTGGGTACAGCTGCAACCCCGAGTGCAGCTGCAGCCCA 1396
DB 2572 AACCTCCATTTCTTGGATGGGTATGAAGTCCATCTGATAAATGGACAGTACAGCCTA 2631

QY 1397 TCGAGCTGCCCGAGAGAGAGTGGACCGTGAACGACATCCAGAAAGCTGTGGGCGCAAGC 1456
DB 2632 TAGTGTGCCAGAAAAAGACAGCTGGACTGTCAATGACATACAGAAAGTGTAGTGGGAAAT 2691

QY 1457 TGAATCTGGCGAGCAGATCTACCCCGGATCAAGGTGGCGCAGCTGTGAAGTGTGTGC 1516
DB 2692 TGAATACCGCAAGTCAAGTATACCCAGGGATTAAGGTAAAGCAATTTATGTAATCTCTTA 2751

QY 1517 GCGGCGCAGGCGCTTACCGACATCTGCCCCCTGACCGAGGAGCGGAGCTGGAGCTGG 1576
DB 2752 GAGGAACCAAGCACTAACAGAGTAAATACCTTAACAGAGAAAGAGAGAGTGAAGCTGG 2811

QY 1577 CCAGAAACCCGAGATCTCTCGCGAGCCGCTGACGCGGTGTACTAGACGCCCGCAGCAAGG 1636
DB 2812 CAGAAAAACAGAGATGTTTAAAGAACCACTACATGAGTGTATTATGACCCATCAAG 2871

QY 1637 ACCTGTGGCGGAGATCCAGAGCAGGCGCCAGCAGTGGACCTACAGATCTACAGG 1696
DB 2872 ACTTAATAGCAGAAATACAGAAAGCAGGGGCAAGGCCAATGGACATATCAAAATTTATCAAG 2931

QY 1697 AGCCCTTCAAGAACCTCAAGACCGCAAGTAGCCCAAGATGCGCACCCCGCCACACCAACG 1756
DB 2932 AGCCATTTAAAAATCTGAAACACAGAAAAATATGCAAGATAGGGGGTGGCCCACTAATG 2991

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:53:38 ; Search time 6233.78 Seconds
(without alignments)
11470.669 Million cell updates/sec

Title: US-09-610-313-32
Perfect score: 2457
Sequence: 1 gtcagcgcaccatggccga.....gggctagcaccgggaattc 2457

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :				
1:	gb_ba.*			AX455916 Sequence
2:	gb_htg.*			AX455915 Sequence
3:	gb_in.*			AX455914 Sequence
4:	gb_em.*			AX427930 Sequence
5:	gb_ov.*			AX427936 Sequence
6:	gb_pat.*			AX427926 Sequence
7:	gb_ph.*			AX427925 Sequence
8:	gb_pl.*			AX427927 Sequence
9:	gb_pr.*			AX427933 Sequence
10:	gb_ro.*			AX427928 Sequence
11:	gb_sts.*			AX427935 Sequence
12:	gb_sy.*			AX427934 Sequence
13:	gb_un.*			AX427932 Sequence
14:	gb_vi.*			AX427921 Sequence
15:	em_ba.*			AF287353 Synthetic
16:	em_fun.*			AX455946 Sequence
17:	em_hum.*			AX455987 Sequence
18:	em_in.*			AF287352 Synthetic
19:	em_mu.*			AX455952 Sequence
20:	em_om.*			AX457088 Sequence
21:	em_or.*			AX306428 Sequence
22:	em_ov.*			AX306429 Sequence
23:	em_pat.*			AX455954 Sequence
24:	em_ph.*			AX427923 Sequence
25:	em_pl.*			AF287355 Synthetic
26:	em_ro.*			AX427937 Sequence
27:	em_sts.*			AX019132 Sequence
28:	em_un.*			AX035453 Sequence
29:	em_vi.*			AX056836 Sequence
30:	em_htg_hum.*			AX283597 Sequence
31:	em_htg_inv.*			AX283607 Sequence
32:	em_htg_other.*			AX035465 Sequence
33:	em_htg_mus.*			AX035464 Sequence
34:	em_htg_pln.*			AX035463 Sequence
35:	em_htg_pln.*			AX283584 Sequence
36:	em_htg_rnd.*			AX455962 Sequence
37:	em_htg_vrt.*			AX455960 Sequence
38:	em_sy.*			AX180752 Sequence
39:	em_htgo_hum.*			AX180747 Sequence
40:	em_htgo_mus.*			AX306430 Sequence
41:	em_htgo_other.*			AX149648 Sequence

ALIGNMENTS

RESULT 1					
AX455916	AX455916	Sequence 32 from Patent WO0204493.	2457 bp	DNA	linear
LOCUS	AX455916	Sequence 32 from Patent WO0204493.			
DEFINITION	AX455916	GI:21714901			
ACCESSION	AX455916.1				
VERSION	AX455916.1				
KEYWORDS		synthetic construct.			
SOURCE		synthetic construct			
ORGANISM		artificial sequences.			
REFERENCE		1			
AUTHORS		zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.			
TITLE		Polynucleotides encoding antigenic hiv type c polypeptides,			
JOURNAL		polypeptides and uses thereof			
		Patent: WO 0204493-A 32 17-JAN-2002;			

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

CHIRON CORPORATION (US) ; University of Stellenbosch (ZA)

Location/Qualifiers

1. -2457

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="PR975YMM"

BASE COUNT 566 a 837 c 754 g 300 t

ORIGIN

Query Match 100.0%; Score 2457; DB 6; Length 2457;

Best Local Similarity 100.0%; Pred. No. 3.1e-248; Indels 0; Gaps 0;

Matches 2457; Conservative 0; Mismatches 0; Gaps 0;

QY 1 GTGAGCGCCACCATGTCGCGAGGCGCATGAGCCAGGCGCCACAGCGCCCAACATCTCTGATGCGAG 60
DB 1 GTGAGCGCCACCATGTCGCGAGGCGCATGAGCCAGGCGCCACAGCGCCCAACATCTCTGATGCGAG 60
QY 61 CGCAGCAACTTCAAGGGCCCCCAAGCGCATCATCAAGTGTTCACATGCGGCGCAAGAGGGC 120
DB 61 CGCAGCAACTTCAAGGGCCCCCAAGCGCATCATCAAGTGTTCACATGCGGCGCAAGAGGGC 120
QY 121 CACATCGCGCAACTGCGCGCGCCCGCCCGCAAGAGGCTGCTGGAAGTGCAGGCAAGGAG 180
DB 121 CACATCGCGCAACTGCGCGCGCCCGCCCGCAAGAGGCTGCTGGAAGTGCAGGCAAGGAG 180
QY 181 GGCACACAGATGAAGGACTGCACCGAGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGCACACAGATGAAGGACTGCACCGAGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTCCCGCAGGCGAAGCGCGCGAGTTCGCCAGCGAGCAGAACCGCGCCCAACAGCCCCACC 300
DB 241 TTCCCGCAGGCGAAGCGCGCGAGTTCGCCAGCGAGCAGAACCGCGCCCAACAGCCCCACC 300
QY 301 AGCGCGAGCTGCAGTGGCGGCGACACCGCCGAGCGAGCGCGCGCGCGCGAGCGCCAG 360
DB 301 AGCGCGAGCTGCAGTGGCGGCGAGAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCCAG 360
QY 361 GGCACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTTGTGAGCATCAAGGTTG 420
DB 361 GGCACCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTTGTGAGCATCAAGGTTG 420
QY 421 GCGCGGCAGATCAAGGAGGCGCTTGGACACCGCGCGCGCGCGCGCGCGCGCGCGAGGAG 480
DB 421 GCGCGGCAGATCAAGGAGGCGCTTGGACACCGCGCGCGCGCGCGCGCGCGCGCGAGGAG 480
QY 481 ATGAGCTGCGCGCAAGTGGAAAGCCCAAGATCATCGCGGCGATCGCGGCGCTTCATCAAG 540
DB 481 ATGAGCTGCGCGCAAGTGGAAAGCCCAAGATCATCGCGGCGATCGCGGCGCTTCATCAAG 540
QY 541 GTGCGCGAGTACGACAGATCTGTGATCGAGATCTGGCGCAAGAGGCCATCGCGCACCGTG 600
DB 541 GTGCGCGAGTACGACAGATCTGTGATCGAGATCTGGCGCAAGAGGCCATCGCGCACCGTG 600
QY 601 CTGATCGGCGCCACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGCTGC 660
DB 601 CTGATCGGCGCCACCCCGTGAACATCATCGGCGCAACATGCTGACCCAGCTGGGCTGC 660
QY 661 ACCCTGAACTTCCCGCATAGCCCGCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATG 720
DB 661 ACCCTGAACTTCCCGCATAGCCCGCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATG 720
QY 721 GAGGCGCCCAAGGTGAAGCAGTGGCCCGCTGACCGAGAGAGATCAAGGCCCTTGACCGCC 780
DB 721 GAGGCGCCCAAGGTGAAGCAGTGGCCCGCTGACCGAGAGAGATCAAGGCCCTTGACCGCC 780
QY 781 ATCTGCGAGGAGATGAGAGGAGGCGCAAGATCACCAAGATCGGCGCCCGGAGACCCCTAC 840
DB 781 ATCTGCGAGGAGATGAGAGGAGGCGCAAGATCACCAAGATCGGCGCCCGGAGACCCCTAC 840
QY 841 AACACCCCGCTGTTCGCCATCAAGAGAGAGGACAGCAGCAAGTGGCGCGCAAGCTGGTGGAC 900
DB 841 AACACCCCGCTGTTCGCCATCAAGAGAGAGGACAGCAGCAAGTGGCGCGCAAGCTGGTGGAC 900

QY 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGTGGGCGATCCCCCAC 960
DB 901 TTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGTGGGCGATCCCCCAC 960
QY 961 CCGCGCGGCTGAAGAAGAAGAGAGCGTGAACCGTGTGAGACGTGGGCGAGCGCTACTTC 1020
DB 961 CCGCGCGGCTGAAGAAGAAGAGAGCGTGAACCGTGTGAGACGTGGGCGAGCGCTACTTC 1020
QY 1021 AGCGTCCCTTGACGAGGAGACTTCCGCAAGTACACCGCTTACACCATCCCGCAGCATCAAC 1080
DB 1021 AGCGTCCCTTGACGAGGAGACTTCCGCAAGTACACCGCTTACACCATCCCGCAGCATCAAC 1080
QY 1081 AACGAGACCCCGGCATCCGCTACCAAGTACAACGTCTGCCCGAGGCGTGAAGGCGAGC 1140
DB 1081 AACGAGACCCCGGCATCCGCTACCAAGTACAACGTCTGCCCGAGGCGTGAAGGCGAGC 1140
QY 1141 CCAGAGATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC 1200
DB 1141 CCAGAGATCTTCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC 1200
QY 1201 GAGATCGTATCTACCAAGGCGCCCTGTACGTGGCGAGCGACCTGGAGATCGGCCAGC 1260
DB 1201 GAGATCGTATCTACCAAGGCGCCCTGTACGTGGCGAGCGACCTGGAGATCGGCCAGC 1260
QY 1261 CGCGCAAGATCGAGGAGTGCAGAGCAGCTCTGCGTGGGCGTTCACACCCCGGAC 1320
DB 1261 CGCGCAAGATCGAGGAGTGCAGAGCAGCTCTGCGTGGGCGTTCACACCCCGGAC 1320
QY 1321 AAGAAGCAGCAGAGGAGCGCCCTTCTGCCCATCGAGCTGCACCCCGCAGAGTGGAC 1380
DB 1321 AAGAAGCAGCAGAGGAGCGCCCTTCTGCCCATCGAGCTGCACCCCGCAGAGTGGAC 1380
QY 1381 GTGAGCCCATCGAGCTGCCGAGAGGAGAGTGGACCGTGAACGACATCCAGAGCTG 1440
DB 1381 GTGAGCCCATCGAGCTGCCGAGAGGAGAGTGGACCGTGAACGACATCCAGAGCTG 1440
QY 1441 GTGGCAAGCTGAACCTGGGCGCAGCATCTACCCGCGCATCAAGGTGCGCCAGCTGTGC 1500
DB 1441 GTGGCAAGCTGAACCTGGGCGCAGCATCTACCCGCGCATCAAGGTGCGCCAGCTGTGC 1500
QY 1501 AAGCTCTGCGCGCGCCAGGCGCTGACGAGATCTGCGCCCTGACCGAGGAGCGCGAG 1560
DB 1501 AAGCTCTGCGCGCGCCAGGCGCTGACGAGATCTGCGCCCTGACCGAGGAGCGCGAG 1560
QY 1561 CTGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCCGTGACGCGCTGTACTACGAC 1620
DB 1561 CTGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCCGTGACGCGCTGTACTACGAC 1620
QY 1621 CCAGCAAGGACCTGGTGGCGGAGATCCAGAGCGGCGCACGACAGTGGACCTACCAG 1680
DB 1621 CCAGCAAGGACCTGGTGGCGGAGATCCAGAGCGGCGCACGACAGTGGACCTACCAG 1680
QY 1681 ATCTACGAGGAGCTTCAAGAAGCTGAACCGCGCAAGTACGCAAGATCGCGCACCGCC 1740
DB 1681 ATCTACGAGGAGCTTCAAGAAGCTGAACCGCGCAAGTACGCAAGATCGCGCACCGCC 1740
QY 1741 CACACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCGAAGATCGCCATGAGAGCATC 1800
DB 1741 CACACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCGAAGATCGCCATGAGAGCATC 1800
QY 1801 GTGATCTGGGCGAAGACCCCGGAGTTCGCGCTGCCCATCCAGAGGAGAGCTGGGAGAC 1860
DB 1801 GTGATCTGGGCGAAGACCCCGGAGTTCGCGCTGCCCATCCAGAGGAGAGCTGGGAGAC 1860
QY 1861 TGGTGGACCCAGTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCG 1920
DB 1861 TGGTGGACCCAGTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGTTCGTGAACACCCCG 1920
QY 1921 CCCTGGTGAAGCTGTGGTACCAAGCTGGAGAGGAGCCCATCATCGCGCGCGGAGACCTTC 1980
DB 1921 CCCTGGTGAAGCTGTGGTACCAAGCTGGAGAGGAGCCCATCATCGCGCGCGGAGACCTTC 1980
QY 1981 TAGTGGACGCGCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCTAGCTGACCGAC 2040


```
Db 1981 TACGTGGACGGCGGCAACCGGAGACCAAGATCGCAAGGCGGTAGCTGACCGAC 2040
Qy 2041 CGGGCCGGCAGAAAGATCGTAGCTGACCGAGACCAACACCAAGAGACCGAGCTGCAG 2100
Db 2041 CGGGCCGGCAGAAAGATCGTAGCTGACCGAGACCAACACCAAGAGACCGAGCTGCAG 2100
Qy 2101 GCCATCCAGCTGGCCCTGACGAGACAGCGGCGAGAGTGAAACATCGTGACCGACAGCCAG 2160
Db 2101 GCCATCCAGCTGGCCCTGACGAGACAGCGGCGAGAGTGAAACATCGTGACCGACAGCCAG 2160
Qy 2161 TACGCCCTGGGCATCATCCAGGCCCGACCCGACCAAGAGCGAGAGCGAGCTGGTGAACCCAG 2220
Db 2161 TACGCCCTGGGCATCATCCAGGCCCGACCCGACCAAGAGCGAGAGCGAGCTGGTGAACCCAG 2220
Qy 2221 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGCCCGCCCAACAAG 2280
Db 2221 ATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGTGCCCGCCCAACAAG 2280
Qy 2281 GGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTG 2340
Db 2281 GGCATCGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCCGCAAGGTGCTG 2340
Qy 2341 TTCTGTGACGCGCATCGATGGCGGCGATCGTGATCTACCAAGTACATGGAGCGCTGTACGTG 2400
Db 2341 TTCTGTGACGCGCATCGATGGCGGCGATCGTGATCTACCAAGTACATGGAGCGCTGTACGTG 2400
Qy 2401 GGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2457
Db 2401 GGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2457

RESULT 2
AX455915
LOCUS AX455915 2463 bp DNA linear PAT 06-JUL-2002
DEFINITION Sequence 31 from Patent WO0204493.
ACCESSION AX455915
VERSION AX455915.1 GI:21714900
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial construct.
REFERENCE
1. zur Megede, J., Barnett, S.W., Engelbrecht, S., and van Rensburg, E.
   Polynucleotides encoding antigenic hiv type c polypeptides,
   polypeptides and uses thereof
   Patent: WO 0204493-A 31 17-JAN-2002;
   CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
    source
    1. .2463
       /organism="synthetic construct"
       /db_xref="taxon:32630"
       /note="PR975YW"
BASE COUNT 567 a 835 c 759 g 302 t
ORIGIN
Query Match 99.2%; Score 2436.2; DB 6; Length 2463;
Best Local Similarity 99.6%; Pred. No. 4.6e-246;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTCGACGCCACCATGGCGGAGGCCATGAGCCAGGCCACCAAGCGCATCTCTGATGCAG 60
Db 1 GTCGACGCCACCATGGCGGAGGCCATGAGCCAGGCCACCAAGCGCATCTCTGATGCAG 60
Qy 61 CGCAGCAACTTCAGGGCCCCCAAGCGCATCATCAAGTGTCTTCACTGGGCAAGAGGGC 120
Db 61 CGCAGCAACTTCAGGGCCCCCAAGCGCATCATCAAGTGTCTTCACTGGGCAAGAGGGC 120
Qy 121 CACATCGCCGCAACTGCCGGCCCCCGCCGCAAGAGGGCTCTGTAAGTGGCGCAAGGAG 180
Db 121 CACATCGCCGCAACTGCCGGCCCCCGCCGCAAGAGGGCTCTGTAAGTGGCGCAAGGAG 180
Qy 181 GGCACCAGATGAAGGACTGCACCGAGCGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCC 240
```

```
Db 181 GGCACCAGATGAAGGACTGCACCGAGCGCCAGGCAACTTCTTCCGCGAGGACCTGGCC 240
Qy 241 TTCCCCCAGGCAAGGCCCGGAGTTCCCGAGCGAGACACCGCCCAACAGCCCCCACC 300
Db 241 TTCCCCCAGGCAAGGCCCGGAGTTCCCGAGCGAGACACCGCCCAACAGCCCCCACC 300
Qy 301 AGCCGCGAGCTGCAGGTGCGGGCGGACAAACCCCGCAGCGAGCGGCGCCGAGCGCCAG 360
Db 301 AGCCGCGAGCTGCAGGTGCGGGCGGACAAACCCCGCAGCGAGCGGCGCCGAGCGCCAG 360
Qy 361 GGCACCTGAACTTCCCGCCAGATCACCTGTGCGAGCGCCCGCTGGTGAAGTCAAGGTG 420
Db 361 GGCACCTGAACTTCCCGCCAGATCACCTGTGCGAGCGCCCGCTGGTGAAGTCAAGGTG 420
Qy 421 GGGGCCAGATCAAGGAGGCCCTGCTGACACCGCGCGGAGCACACCTGCTGGAGGAG 480
Db 421 GGGGCCAGATCAAGGAGGCCCTGCTGACACCGCGCGGAGCACACCTGCTGGAGGAG 480
Qy 481 ATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGCGGCGATCGGGGCTTCATCAAG 540
Db 481 ATGAGCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGCGATCGGGGCTTCATCAAG 540
Qy 541 GTGCGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600
Db 541 GTGCGCCAGTACGACACAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTG 600
Qy 601 CTGATCGGCCCGCCCGGCGTGAACATCATCGGCGCAACATGCTGACCGAGCTGGGCTGC 660
Db 601 CTGATCGGCCCGCCCGGCGTGAACATCATCGGCGCAACATGCTGACCGAGCTGGGCTGC 660
Qy 661 ACCCTGAACTTCCCGCATCAGCCCCATCGAGACCGCGTGCAGCTGAAGCCGCGCATG 720
Db 661 ACCCTGAACTTCCCGCATCAGCCCCATCGAGACCGCGTGCAGCTGAAGCCGCGCATG 720
Qy 721 GACGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTTGACCGCC 780
Db 721 GACGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTTGACCGCC 780
Qy 781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAGATCGGCGCCGAGAACCCCTAC 840
Db 781 ATCTGCGAGGAGATGGAGAAGGAGGGCAAGATCACCAGATCGGCGCCGAGAACCCCTAC 840
Qy 841 AACACCCCGCTGTTCCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGAC 900
Db 841 AACACCCCGCTGTTCCGCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGTGGAC 900
Qy 901 TTCCGCGAGCTGAACAAGCGCACCCAGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
Db 901 TTCCGCGAGCTGAACAAGCGCACCCAGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
Qy 961 CCGCGCGCCTGAGAAGAAGAGCGCTGACCGTGTGGAGCGCGCTGAGCGCTGGGCGAGCTTTC 1020
Db 961 CCGCGCGCCTGAGAAGAAGAGCGCTGACCGTGTGGAGCGCGCTGAGCGCTGGGCGAGCTTTC 1020
Qy 1021 AGCGTGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAAC 1080
Db 1021 AGCGTGCCCTGGAGGAGACTTCCGCAAGTACACCGCTTCAACATCCCGAGCATCAAC 1080
Qy 1081 AACGAGACCCCGCGCATCCGCTACAGTACAACTGTCTGCCCGCGAGGCTGGAAGGCGAGC 1140
Db 1081 AACGAGACCCCGCGCATCCGCTACAGTACAACTGTCTGCCCGCGAGGCTGGAAGGCGAGC 1140
Qy 1141 CCGAGCATCTTCCAGAGCAGATGACCAAGATCTTGGAGCGCTTCCGCGCGCGCAACCC 1200
Db 1141 CCGAGCATCTTCCAGAGCAGATGACCAAGATCTTGGAGCGCTTCCGCGCGCGCAACCC 1200
Qy 1201 GAGATCGTGATCTACCAAGCGCCCTGCTGACGTGGCAGCGACCTGAGATCGGCGAGCAGC 1260
Db 1201 GAGATCGTGATCTACCAAGCGCCCTGCTGACGTGGCAGCGACCTGAGATCGGCGAGCAGC 1260
Qy 1261 CGCGCCAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCAACCCCGCAC 1320
Db 1261 CGCGCCAGATCGAGGAGCTGCGCAAGCACCTGCTGCGCTGGGGCTTCACCAACCCCGCAC 1320
```

Db 1261 CGGCCAAGATCAGGAGCTGCCAAGACACCTGCTGGCTGGGGCTTACCACCCCGGAC 1320
QY 1321 AAGAAGCACAGAGAGAGCCCTTCTCCCAT-----CGAGCTGCACCCCGACAAG 1374
Db 1321 AAGAAGCACAGAGAGAGCCCTTCTCCCAT-----CGAGCTGCACCCCGACAAG 1380
QY 1375 TGGACCGTGAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACGACATCCAG 1434
Db 1381 TGGACCGTGAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACGACATCCAG 1440
QY 1435 AAGCTGTGGGCAAGCTGAACCTGGCGAGCAGCATATACCCCGCATCAAGGTGCCCGAG 1494
Db 1441 AAGCTGTGGGCAAGCTGAACCTGGCGAGCAGCATATACCCCGCATCAAGGTGCCCGAG 1500
QY 1495 CTGTGAAGCTGTGGCGGCGCAAGGCGCTGACCGACATCTGCGCCCTGACCGAGGAG 1554
Db 1501 CTGTGAAGCTGTGGCGGCGCAAGGCGCTGACCGACATCTGCGCCCTGACCGAGGAG 1560
QY 1555 GCCGAGCTGAGCTGGCGGAGAACCGCGAGATCCTGGCGAGCCCGTGCACGCGGTGAC 1614
Db 1561 GCCGAGCTGAGCTGGCGGAGAACCGCGAGATCCTGGCGAGCCCGTGCACGCGGTGAC 1620
QY 1615 TACGACCCAGCAAGACCTGTGGCGGAGATCCAGAAGCAGGCGCACGACCATGGACC 1674
Db 1621 TACGACCCAGCAAGACCTGTGGCGGAGATCCAGAAGCAGGCGCACGACCATGGACC 1680
QY 1675 TACGACATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1734
Db 1681 TACGACATCTACGAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGC 1740
QY 1735 ACCGCCCCACCAAGACGTGAAGAGCTGACGAGGCGCGTGAGAGATCGCCATGGAG 1794
Db 1741 ACCGCCCCACCAAGACGTGAAGAGCTGACGAGGCGCGTGAGAGATCGCCATGGAG 1800
QY 1795 AGCATGCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAGAGACCTGG 1854
Db 1801 AGCATGCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAGAGACCTGG 1860
QY 1855 GAGACCTGTGTGACCGACTACTGGCAGCGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1914
Db 1861 GAGACCTGTGTGACCGACTACTGGCAGCGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
QY 1915 ACCGCCCCCTGTGAGCTGTGGTACGAGCTGGAGAGAGCCCATCATCGCGCGCGAG 1974
Db 1921 ACCGCCCCCTGTGAGCTGTGGTACGAGCTGGAGAGAGCCCATCATCGCGCGCGAG 1980
QY 1975 ACCTTCTACGTGACGCGCGCCCAACCGGAGACCAAGATCGCAAGCGCGCTACGTG 2034
Db 1981 ACCTTCTACGTGACGCGCGCCCAACCGGAGACCAAGATCGCAAGCGCGCTACGTG 2040
QY 2035 ACCGACCGGGCGGAGAGATCGTGAGCTGACCGAGACCAACCAAGAGACCGAG 2094
Db 2041 ACCGACCGGGCGGAGAGATCGTGAGCTGACCGAGACCAACCAAGAGACCGAG 2100
QY 2095 CTGAGGCCATCCAGCTTGGCCCTTGCAAGACAGCGGAGAGGTGAACATCGTGACCGAG 2154
Db 2101 CTGAGGCCATCCAGCTTGGCCCTTGCAAGACAGCGGAGAGGTGAACATCGTGACCGAG 2160
QY 2155 AGCAGTACGCCCTTGGCATCATCCAGGCCACCGCGACAGAGCGAGCGAGCTGGTG 2214
Db 2161 AGCAGTACGCCCTTGGCATCATCCAGGCCACCGCGACAGAGCGAGCGAGCTGGTG 2220
QY 2215 AACGAGATCATCGAGAGCTGTGAAGAGGAGAGTGTACCTGAGCTGGGTGCCCGCC 2274
Db 2221 AACGAGATCATCGAGAGCTGTGAAGAGGAGAGTGTACCTGAGCTGGGTGCCCGCC 2280
QY 2275 CACAAGGCATCGGCGGCAACGAGCAGATGACAAAGCTGGTGAAGAGGCGCATCCGCAAG 2334
Db 2281 CACAAGGCATCGGCGGCAACGAGCAGATGACAAAGCTGGTGAAGAGGCGCATCCGCAAG 2340
QY 2335 GTGCTGTTCTGACCGCATCGATGCGGCGATCGTATCTACCAAGTACATGGACGACCTG 2394
Db 2341 GTGCTGTTCTGACCGCATCGATGCGGCGATCGTATCTACCAAGTACATGGACGACCTG 2400

QY 2395 TACGTGGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2454
Db 2401 TACGTGGCAGCGCGGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
QY 2455 TTC 2457
Db 2461 TTC 2463
RESULT 3
AX455914 AX455914 2469 bp DNA linear PAT 06-JUL-2002
LOCUS Sequence 30 from Patent WO0204493.
DEFINITION AX455914
ACCESSION AX455914
VERSION AX455914.1 GI:21714899
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS zur Megede, J., Barnett, S.W., Engelbrecht, S. and van Rensburg, E.
TITLE Polynucleotides encoding antigenic hiv type c polypeptides,
polypeptides and uses thereof
JOURNAL Patent: WO 0204493-A 30 17-JAN-2002;
CHIRON CORPORATION (US); University of Stellenbosch (ZA)
FEATURES
Location/Qualifiers
1..2469
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="PR975(+)"
BASE COUNT 571 a 833 c 761 g 304 t
ORIGIN
Query Match 98.3%; Score 2415.4; DB 6; Length 2469;
Best Local Similarity 99.3%; Pred. No. 6.9e-244;
Matches 2451; Conservative 0; Mismatches 6; Indels 12; Gaps 2;
QY 1 GTGACGCCACCATGGCGAGGCGCATGAGCCAGGCCACCGCCACACATCTCTGATGAG 60
Db 1 GTGACGCCACCATGGCGAGGCGCATGAGCCAGGCCACCGCCACATCTCTGATGAG 60
QY 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGTTCACATCGCGCAAGAGGGC 120
Db 61 CGCAGCAACTTCAAGGGCCCAAGCGCATCATCAAGTGTTCACATCGCGCAAGAGGGC 120
QY 121 CACATCGCCCGCAACTGCGCGGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGAG 180
Db 121 CACATCGCCCGCAACTGCGCGGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGGCAAGAG 180
QY 181 GGCACCCAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTCCGCGAGGACCTGGCC 240
Db 181 GGCACCCAGATGAAGGACTGCACCGAGCGCCAGCCAACTTCTCCGCGAGGACCTGGCC 240
QY 241 TTCCCCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGCAACCGCCGCAACAGCCCCACC 300
Db 241 TTCCCCCAGGGCAAGGCCCGCGAGTTCCCGAGCGAGCAACCGCCGCAACAGCCCCACC 300
QY 301 AGCGCGAGCTGAGGTGCGCGGGGACAAACCCCGAGCGAGCGCGCGCGCGCGCGAG 360
Db 301 AGCGCGAGCTGAGGTGCGCGGGGACAAACCCCGAGCGAGCGCGCGCGCGCGCGAG 360
QY 361 GGCACCTGAACCTTCCCGCCAGATCACCTGTGCGAGCGCCCTTGGTGAAGTCAAGGTG 420
Db 361 GGCACCTGAACCTTCCCGCCAGATCACCTGTGCGAGCGCCCTTGGTGAAGTCAAGGTG 420
QY 421 GGCAGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGAG 480
Db 421 GGCAGCCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGAG 480
QY 481 ATGAGCCTGCCCCGCGCAAGTGGAGGCCCAAGATGATGCGCGCGCATCGCGCGCTTCATCAAG 540
Db 481 ATGAGCCTGCCCCGCGCAAGTGGAGGCCCAAGATGATGCGCGCGCATCGCGCGCTTCATCAAG 540

QY 541 GTGCGCAGTACACAGATCCCTGATCGAGATCTGGGCAAGAGGCCATCGCACCGTG 600
Db 541 GTGCGCAGTACACAGATCCCTGATCGAGATCTGGGCAAGAGGCCATCGCACCGTG 600
QY 601 CTGATCGGCCCCACCCCGGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGTGC 660
Db 601 CTGATCGGCCCCACCCCGGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGTGC 660
QY 661 ACCCTGAACCTCCCGATCAGCCCATCAGCAGACCGTCCCGTGAAGCTGAAGCCCGCATG 720
Db 661 ACCCTGAACCTCCCGATCAGCCCATCAGCAGACCGTCCCGTGAAGCTGAAGCCCGCATG 720
QY 721 GAGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGAGAGATCAAGGGCCCTGACCGCC 780
Db 721 GAGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGAGAGATCAAGGGCCCTGACCGCC 780
QY 781 ATCTGAGAGATGAGAGAGAGAGAGATCAACCAAGATCGGCCCGCGAGACCCCTAC 840
Db 781 ATCTGAGAGATGAGAGAGAGAGAGATCAACCAAGATCGGCCCGCGAGACCCCTAC 840
QY 841 AACACCCCGCTGTTCGCCATCAAGAGAGAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCGCTGTTCGCCATCAAGAGAGAGAGACACCAAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGCGAGCTGAACAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
Db 901 TTCGCGAGCTGAACAGCGACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCCAC 960
QY 961 CCCGCGGCTGAAGAGAGAGAGAGAGCTGACCGTGTGGAGCTGGCGACGCGCTACTTC 1020
Db 961 CCCGCGGCTGAAGAGAGAGAGAGAGCTGACCGTGTGGAGCTGGCGACGCGCTACTTC 1020
QY 1021 AGCGTCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACATCCCGAGCATCAAC 1080
Db 1021 AGCGTCCCTGGAGAGGACTTCCGCAAGTACACCGCTTCCACATCCCGAGCATCAAC 1080
QY 1081 AACGAGACCCCGGCTACCGCTACCAAGTGTGTCGCCCGAGGCTGGAGGGCAGC 1140
Db 1081 AACGAGACCCCGGCTACCGCTACCAAGTGTGTCGCCCGAGGCTGGAGGGCAGC 1140
QY 1141 CCAGCATCTTCAGAGCAGCATGACCAAGATCTGAGAGCCCTTCCGCGCGCGCAACCC 1200
Db 1141 CCAGCATCTTCAGAGCAGCATGACCAAGATCTGAGAGCCCTTCCGCGCGCGCAACCC 1200
QY 1201 GAGATCGTATCTACCA- - - - -GGCCCGCTGTACCTGGGCGAGCAGCTGGAGTCGC 1254
Db 1201 GAGATCGTATCTACCAAGTACATGGAGCAGCTGTACCTGGGCGAGCAGCTGGAGTCGC 1254
QY 1255 CAGCACCGCCCAAGATCGAGGAGCTCGCAAGCAGCTGCTGGCTGGGCTTCAACACC 1314
Db 1255 CAGCACCGCCCAAGATCGAGGAGCTCGCAAGCAGCTGCTGGCTGGGCTTCAACACC 1314
QY 1315 CCGGACAAGAGCAGCAGAGAGAGCCCTTCTGTCGCCAT- - - - -CGAGCTGCACCC 1368
Db 1315 CCGGACAAGAGCAGCAGAGAGAGCCCTTCTGTCGCCAT- - - - -CGAGCTGCACCC 1368
QY 1369 GACAAGTGGAGCTGACGCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACGAC 1428
Db 1369 GACAAGTGGAGCTGACGCCATCGAGCTGCCGAGAGAGAGCTGGACCTGGAACGAC 1428
QY 1429 ATCCAGAGCTGGTGGCAAGCTGAACCTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG 1488
Db 1429 ATCCAGAGCTGGTGGCAAGCTGAACCTGGGCCAGCCAGATCTACCCCGGCATCAAGGTG 1488
QY 1489 CGCAGCTGTGCAAGCTGTGCGGGGCGCAAGCCCTGACCGACATCTGTCGCCCTGACC 1548
Db 1489 CGCAGCTGTGCAAGCTGTGCGGGGCGCAAGCCCTGACCGACATCTGTCGCCCTGACC 1548
QY 1549 GAGGAGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTTGGCGAGCCGCTGCACGGC 1608
Db 1549 GAGGAGCCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTTGGCGAGCCGCTGCACGGC 1608

QY 1609 GTGTACTACACCCACCAAGGACCTGTGTGGCCGAGATCCAGAAAGAGGCCACGACAC 1668
Db 1609 GTGTACTACACCCACCAAGGACCTGTGTGGCCGAGATCCAGAAAGAGGCCACGACAC 1668
QY 1669 TGGACCTTACAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCAAG 1728
Db 1669 TGGACCTTACAGATCTTACCAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCAAG 1728
QY 1729 ATGCGCACCCGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCGCC 1788
Db 1729 ATGCGCACCCGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCGCC 1788
QY 1741 ATGCGCACCCGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCGCC 1800
Db 1741 ATGCGCACCCGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGCTGCAAGAGATCGCC 1800
QY 1789 ATGAGAGATCTGTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAGAG 1848
Db 1789 ATGAGAGATCTGTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAGAG 1848
QY 1801 ATGAGAGATCTGTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAGAG 1860
Db 1801 ATGAGAGATCTGTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAAGAG 1860
QY 1849 ACCTGGAGACCTGTGTGAGCCGACTACTGGAGGCCACCTGGATCCCGAGTGGGAGTTC 1908
Db 1849 ACCTGGAGACCTGTGTGAGCCGACTACTGGAGGCCACCTGGATCCCGAGTGGGAGTTC 1908
QY 1909 GTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGC 1968
Db 1909 GTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGC 1968
QY 1921 GTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGC 1980
Db 1921 GTGAACACCCCGCTGTGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGC 1980
QY 1969 GCCGAGACCTTCTACGTGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2028
Db 1969 GCCGAGACCTTCTACGTGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2028
QY 1981 GCCGAGACCTTCTACGTGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2040
Db 1981 GCCGAGACCTTCTACGTGAGCGCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGC 2040
QY 2029 TACGTGACCGACCGCGCGCGCGAGAGATCGTGAAGCTGACCGTACCGAGACCAACGAGAG 2088
Db 2029 TACGTGACCGACCGCGCGCGCGAGAGATCGTGAAGCTGACCGTACCGAGACCAACGAGAG 2088
QY 2041 TACGTGACCGACCGCGCGCGCGAGAGATCGTGAAGCTGACCGTACCGAGACCAACGAGAG 2100
Db 2041 TACGTGACCGACCGCGCGCGCGAGAGATCGTGAAGCTGACCGTACCGAGACCAACGAGAG 2100
QY 2089 ACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGTTGAACATCGTG 2148
Db 2089 ACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGTTGAACATCGTG 2148
QY 2101 ACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGTTGAACATCGTG 2160
Db 2101 ACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGGTTGAACATCGTG 2160
QY 2149 ACCGAGCAGTACGCGCTGGGCATCATCAGGCCCGAGCCCGAGAGCGAGAGCGAG 2208
Db 2149 ACCGAGCAGTACGCGCTGGGCATCATCAGGCCCGAGCCCGAGAGCGAGAGCGAG 2208
QY 2161 ACCGAGCAGTACGCGCTGGGCATCATCAGGCCCGAGCCCGAGAGCGAGAGCGAG 2220
Db 2161 ACCGAGCAGTACGCGCTGGGCATCATCAGGCCCGAGCCCGAGAGCGAGAGCGAG 2220
QY 2209 CTGCTGACCGACGATCATCGAGCAGCTGATCAAGAGGAGAGGTTGACCTGAGCTGGGTG 2268
Db 2209 CTGCTGACCGACGATCATCGAGCAGCTGATCAAGAGGAGAGGTTGACCTGAGCTGGGTG 2268
QY 2221 CTGCTGACCGACGATCATCGAGCAGCTGATCAAGAGGAGAGGTTGACCTGAGCTGGGTG 2280
Db 2221 CTGCTGACCGACGATCATCGAGCAGCTGATCAAGAGGAGAGGTTGACCTGAGCTGGGTG 2280
QY 2269 CCGCCCAAGAGGCGATCGCGCGCAACGAGCAGATCGAAGAGCTGGTGAAGAGGCGCATC 2328
Db 2269 CCGCCCAAGAGGCGATCGCGCGCAACGAGCAGATCGAAGAGCTGGTGAAGAGGCGCATC 2328
QY 2281 CCGCCCAAGAGGCGATCGCGCGCAACGAGCAGATCGAAGAGCTGGTGAAGAGGCGATC 2340
Db 2281 CCGCCCAAGAGGCGATCGCGCGCAACGAGCAGATCGAAGAGCTGGTGAAGAGGCGATC 2340
QY 2329 CGCAAGTGTGTTCTTGGAGCGCATCGATGGCGCATCTGATGATCTACCAAGTACATGAC 2388
Db 2329 CGCAAGTGTGTTCTTGGAGCGCATCGATGGCGCATCTGATGATCTACCAAGTACATGAC 2388
QY 2341 CGCAAGTGTGTTCTTGGAGCGCATCGATGGCGCATCTGATGATCTACCAAGTACATGAC 2400
Db 2341 CGCAAGTGTGTTCTTGGAGCGCATCGATGGCGCATCTGATGATCTACCAAGTACATGAC 2400
QY 2389 GACCTGTACCTGGGCGAGCGCGCGCTAGGATCGATTAAGAGCTTCCCGGGCTAGACACC 2448
Db 2389 GACCTGTACCTGGGCGAGCGCGCGCTAGGATCGATTAAGAGCTTCCCGGGCTAGACACC 2448
QY 2401 GACCTGTACCTGGGCGAGCGCGCGCTAGGATCGATTAAGAGCTTCCCGGGCTAGACACC 2460
Db 2401 GACCTGTACCTGGGCGAGCGCGCGCTAGGATCGATTAAGAGCTTCCCGGGCTAGACACC 2460
QY 2449 GGTGAATTC 2457
Db 2449 GGTGAATTC 2457
QY 2461 GGTGAATTC 2469
Db 2461 GGTGAATTC 2469

RESULT 4

AX427930

LOCUS

AX427930

DEFINITION

Sequence 168 from Patent WO0232943.

ACCESSION

AX427930

VERSION

AX427930.1

KEYWORDS

GI:21538017

SOURCE

synthetic construct.

ORGANISM

synthetic construct.

artificial sequences.

REFERENCE

1

AUTHORS

Huang, Y. and Nabel, G.J.

PAT 20-JUN-2002

linear

9166 bp DNA

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 168 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES Location/Qualifiers
source 1..9166
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2227 a 2662 c 2513 g 1764 t
ORIGIN

Query Match 81.4%; Score 2000.8; DB 6; Length 9166;
Best Local Similarity 91.3%; Pred. No. 1e-200;
Matches 2185; Conservative 0; Mismatches 182; Indels 25; Gaps 5;
QY 14 TGGCCGAGGCATGAGCGAGCCACCGACCACTCTGTGATGCGGAGCAACCTTCA 73
DB 2961 TGGCCGAGGCATGAGCAAGTGNACAAACCAACATCATGTGACGCGAGCAACTGCA 3020
QY 74 AGGCCCCAAAGCGCATCATCAAGTGTCTTCAACTGCGCAAGGAGGCGCACATCGCCGCA 133
DB 3021 AGGCCCCAAAGCGCACCATCAAGTGTCTTCAACTGCGCAAGGAGGCGCACCTGCGCCGCA 3080
QY 134 ACTGCGCGCCCGCCGCAAGAGGGCTGCTGGAGTCCGCAAGGAGGCGCACCATGTA 193
DB 3081 ACTGCGCGCCCGCCGCAAGAGGGCTGCTGGAGTCCGCAAGGAGGCGCACCATGTA 3140
QY 194 AGGACTCACCGAGCGCGGACCACTTCTTCCGCGAGGACCTTGGCTTCCCGCAGGGCA 253
DB 3141 AGGACTCACCGAGCGACAGGCTAA-TTTTATAGGAGATCTGGCCCTTCCCAAGGGA 3199
QY 254 AGGCGCGCGAGTTCCCGAGCGAGAGCAACCGCGCCACACAGCCCGGAGGCTGC 313
DB 3200 AGGCGAGGAATTTCTTTCAGAGCAGACAGAGCCACACAGCCCGCAGAGAGAGCTC 3259
QY 314 AGGTGCGCGG-----CGAACACCCCGCAGAGGCGCGGCGGAGCGCCAGGGCA--- 364
DB 3260 AGGTGCGCGGAGAGAGCAACTTCCCTCTCAGAAGCAGGAGCGGATAGCAAGGAAGT 3319
QY 365 ---CCCTGAATTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGTGG 421
DB 3320 TATCCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCGTCAACAATAAGATAG 3379
QY 422 CGCGCCAGATCAAGAGGCGCCTGCTGGACACCGCGCGCCGACACACCGTCTGGAGAGA 481
DB 3380 GGGCCAGCTGAAGAGGCGCCTTTAGACACCGCGCGCCGACACACCGTCTGGAGGAGA 3439
QY 482 TGAGCCTGCGCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCAAG 541
DB 3440 TGAACCTGCGCGCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGGCTTTCATCAAG 3499
QY 542 TGCGCCAGTACGACCATCTGATCGAGATCTGCGGCAAGAAGGCCATCGGACCGCTGC 601
DB 3500 TGGCCAGTACGACCATCTGATCGAGATCTGCGGCCACAAGGCCATCGGACCGCTGC 3559
QY 602 TGATCGGCCCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGCA 661
DB 3560 TGGTGGGCCCCACCGCGTGAACATCATCGCGCGCAACCTGCTGCCACATCGGCTGCA 3619
QY 662 CCCTGACTTCCCGCATCAGCCCATCGAGACCGTGGCCGTGGAAGCTGGAAGCCCGCATGG 721
DB 3620 CCCTGACTTCCCGCATCAGCCCATCGAGACCGTGGCCGTGGAAGCTGGAAGCCCGCATGG 3679
QY 722 ACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACGCCCA 781
DB 3680 ACGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGGTGAGA 3739
QY 782 TCTGCGAGGAGATGGAGAGGAGGCAAGATCATCCAAGATCGGCCCGCGAGAACCCCTTACA 841
DB 3740 TCTGCAACCGAGATGGAGAGGAGGCAAGATCATCCAAGATCGGCCCGCGAGAACCCCTTACA 3799
QY 842 ACACCCCGCTGTTGCGCATCAAGAAGAGGACAGCAACCAAGTGGCGCAAGCTGGTGA 901

DB 3800 ACACCCCGTGTTCGCCATCAAGAAGAGGACACCAAGTGGCGCAAGCTGGTGGACT 3859
QY 902 TCCCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGATCCCCCACC 961
DB 3860 TCCCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGACGTGGGATCCCCCACC 3919
QY 962 CCGCGCGCTGAAGAAGAAGAGCGTGCCTGGAGTGGGCGGACGCTTACTTCA 1021
DB 3920 CCGCGCGCTGAAGAAGAAGAGCGTGCCTGGAGTGGGCGGACGCTTACTTCA 3979
QY 1022 GCGTGCCCTGGAGAGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCAACA 1081
DB 3980 GCGTGCCCTGGACAAGGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCAACA 4039
QY 1082 AGGAGACCCCGGATCGCTTACCAGTACACGCTGCTGCCCGAGGCTGGAGGCGAGCC 1141
DB 4040 ACGAGACCCCGGATCGCTTACCAGTACACGCTGCTGCCCGAGGCTGGAGGCGAGCC 4099
QY 1142 CAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCGCCGCAACCCG 1201
DB 4100 CCGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAACCCG 4159
QY 1202 AGATGATGATCTACCA-----GGCCCCCTGTAGCTGGGCGAGCCTGGAGATCGGCC 1255
DB 4160 ACATGCTGATCTACCAAGTACATGACCAACCTGTACGTGGCGAGCAGCTGGAGATCGGCC 4219
QY 1256 AGACCCCGCAAGATCGAGAGCTGCGCAAGCAGCTGCTGCGTGGGCTTTCACCAACC 1315
DB 4220 AGCACCCAGCAGATCGAGAGCTGCGCAGCAGCTGCTGCGTGGGCTTTCACCAACC 4279
QY 1316 CCGCAAGAAGACACCAAGAGGAGCCCTTCTCTGCCAT-----CGAGGTGCACCCCG 1369
DB 4280 CCGCAAGAAGACACCAAGAGGAGCCCTTCTCTGGATGGGTACGAGTGCACCCCG 4339
QY 1370 ARAAGTGGACCTGTCACCCCATCGAGCTGCCGAGAGAGAGAGCTGACCGTGAACGACA 1429
DB 4340 ACAAGTGGACCTGTCACCCCATCGTGTGCCGAGAGAGAGCTGACCGTGAACGACA 4399
QY 1430 TCAGAAAGCTGGTGGCAAGCTGAACCTGGCCAGCAGATCTACCCCGCATCAAGTGC 1489
DB 4400 TCAGAAAGCTGGTGGCAAGCTGAACCTGGCCAGCAGATCTACCCCGCATCAAGTGC 4459
QY 1490 GCAGCTGTCAAGCTGCTCGCGCGCCCAAGGCCCTGACCGACATCTGCGCCCTGACC 1549
DB 4460 GCAGCTGTCAAGCTGCTCGCGCGCCCAAGGCCCTGACCGAGGTGCTGCCCTGACC 4519
QY 1550 AGAGGCCGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGGCGAGCCCGTGCACGGCG 1609
DB 4520 AGAGGCCGAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGAAGGAGCCCGTGCACGGCG 4579
QY 1610 TGTACTACGACCCAGCAAGGACTGTGTGGCGAAGACCCCAAGTTCGCCCTGCCATCCAGAGGAGA 1669
DB 4580 TGTACTACGACCCAGCAAGGACTGTGTGGCGAAGACCCCAAGTTCAGCTGCCATCCAGAGGAGA 4639
QY 1670 GGACCTACAGATCTACAGGAGCCCTTCAAGACCTGAAGACCGGCAAGTACGCCAAGA 1729
DB 4640 GGACCTACAGATCTACAGGAGCCCTTCAAGACCTGAAGACCGGCAAGTACGCCGCA 4699
QY 1730 TGGCACCCCGCCACCAAGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAGATGCCCA 1789
DB 4700 TGAAGGCGCCCGCCACCAAGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAGATGCCCA 4759
QY 1790 TGGAGAGCATCTGTGGGCAAGACCCCAAGTTCGCCCTGCCATCCAGAGGAGA 1849
DB 4760 CCGAGAGCATCTGTGGGCAAGACCCCAAGTTCAGCTGCCATCCAGAGGAGA 4819
QY 1850 CTTGGAGAGCTGTGGAGCAGCTACTTGGCAGGCGACCTGGATCCCCAGTGGAGTTCG 1909
DB 4820 CTTGGAGAGCTGTGGAGCAGCTACTTGGCAGGCGACCTGGATCCCCAGTGGAGTTCG 4879
QY 1910 TGAACACCCCGCCCGCTGTGAACTGTGGTACCAGCTGGAGAGGAGGCCCATCATCGCGG 1969

Db 4880 TGAACACCCCCCTTCTGTTGAAGCTGTGGTACCAAGTGGAGAGAGGCCCATCATCGGCG 4939

QY 1970 CCGAGACCTTCTACCTGGACGGCGCCGCAACCGCGAGACCAAGATCGGCAAGCGCGCT 2029

Db 4940 CCGAGACCTTCTACCTGGACGGCGCCGCAACCGCGAGACCAAGTGGCAAGCGCGCT 4999

QY 2030 AGCTGACCGACCGGGCGCGGAGAGATCGTGAAGCTGACCGAGACCAACCAAGAGA 2089

Db 5000 AGCTGACCGACCGGGCGCGGAGAGTGGTCCCTGACCGACCAACCAAGAGA 5059

QY 2090 CCGAGCTGACGGCCATCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCGTGA 2149

Db 5060 CCGAGCTGACGGCCATCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCGTGA 5119

QY 2150 CCGAGCGACGTACCGCTTGGGCATCTCCAGGCGCCAGCGGCAAGAGCGAGCGAGC 2209

Db 5120 CCGAGCGACGTACCGCTTGGGCATCTCCAGGCGCCAGCGGCAAGAGCGAGCGAGC 5179

QY 2210 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGC 2269

Db 5180 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGTGC 5239

QY 2270 CGGCCCCAAGGGATCGGCGCACGACGACATCGACAAGCTGTGAGCAAGGCGATCC 2329

Db 5240 CGGCCCCAAGGGATCGGCGCACGACGACGATCGACAAGCTGTGAGCAAGGCGATCC 5299

QY 2330 GCAAGGTGCTTCTCGACGCGCATCGATGGCGGCTCGTGTATCTACCACTA 2381

Db 5300 GCAAGGTGCTTCTCGACGCGCATCGACAGGCGCCAGGAGGAGGACCAAGAA 5351

RESULT 5

AX427936

LOCUS AX427936 9788 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 174 from Patent WO0232943.

ACCESSION AX427936

VERSION AX427936.1 GI:21538023

KEYWORDS synthetic construct.

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Huang,Y. and Nabel,G.J.

TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization

JOURNAL Patent: WO 0232943-A 174 25-APR-2002;

GOVERNMENT OF THE UNITED STATES (US)

FEATURES

Location/Qualifiers

1..9788

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2377 a 2817 c 2695 g 1899 t

ORIGIN

Query Match 81.1%; Score 1992.4; DB 6; Length 9788;

Best Local Similarity 91.2%; Pred. No. 7.8e-200;

Matches 2181; Conservative 0; Mismatches 181; Indels 30; Gaps 5;

QY 14 TGGCCGAGGCCATGAGCAGGCGCCAGCCAGCCACATCTCTGATGAGCGGCAACTTCA 73

Db 2961 TGGCCGAGGCCATGAGCAGGCGCCAGCCAGCCACATCTCTGATGAGCGGCAACTTCA 3020

QY 74 AGGGCCCCAAGCGCATCATCAAGTGTCTTCACTGGCGGCAAGGCGCCACATCGCCGCA 133

Db 3021 AGGGCCCCAAGCGCACCATCAAGTGTCTTCACTGGCGGCAAGGCGCCACATCGCCGCA 3080

QY 134 ACTCGCGCGCCCCCGCAAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCGCCACAGATGA 193

Db 3081 ACTCGCGCGCCCCCGCAAGAGGGCTGTCTGGAAGTGGCGCAAGGAGGCGCCACAGATGA 3140

QY 194 AGGACTGACCGGCGCCAGGCGCAACTTCTTCCGCGAGGAGCTGGCGCTTCCCGAGGCA 253

Db 3141 AGGACTGACCGGAGCGACAGGCTAA-----TAGGGAAGATCTGGCTTCCCAAGGGA 3194

QY 254 AGGCGCGGAGTTCCTCCACGAGCAGAACCGCGCCCAACAGAGCCCCACGAGCGGAGCTGC 313

Db 3195 AGGCGCGGAGTTCCTCCACGAGCAGAACCGCGCCCAACAGAGCCCCACGAGAGAGCTTC 3254

QY 314 AGGTGCGGG-----CGACAACCCCGCAGCGAGCGCGCGCGCGCGCCAGGCA--- 364

Db 3255 AGTTTGGGAGAGACACAACTCCCTCTCAGAGCAGAGCGCGGATAGACAAGGAATG 3314

QY 365 ---CCCTGAACCTTCCCGCCAGATCACCTGTGGCAGCGCCCCCTGTGTGAGCATCAAGTGG 421

Db 3315 TATCCCTTTAGCTTCCTCAGATCACTTCTTGGCAGGACCCCTCGTCAATAAAGATAG 3374

QY 422 GCGGCAGATCAAGGAGGCGCTGTGGACACCGCGCGCCGACACACCGTGTGGAGGAGA 481

Db 3375 GGGGCGAGCTGAAGGAGGCGCTTCTAGACCGCGCGCGCGACACCGTGTGGAGGAGA 3434

QY 482 TGAGCTTGGCGGCAAGTGGAGCGCAAGATGATCGCGGCGATCGCGGCTTTCATCAAGG 541

Db 3435 TGAACCTGCCCGCGCTGGAAGCGCAAGATGATCGCGGCGATCGCGGCTTTCATCAAGG 3494

QY 542 TCGGCGAGTACGACGAGATCTGTGATCGAGATCTGGCGCAAGAGCGCATCGCACCGTGC 601

Db 3495 TGGGCGAGTACGACGAGATCTGTGATCGAGATCTGGCGCAAGAGCGCATCGCACCGTGC 3554

QY 602 TGATCGGCGCCACCGCTGAAACATCATCGGCGCAAGATGATCGCACCGTGTGGGTGCA 661

Db 3555 TGGTGGGCGCCACCGCTGAAACATCATCGGCGCAAGATGATCGCACCGTGTGGGTGCA 3614

QY 662 CCCTGAACCTTCCCGCATCGAGCGCTGAGACCGTGGCGGCTGAAGCTGAAGCGCGCATGG 721

Db 3615 CCCTGAACCTTCCCGCATCGAGCGCTGAGACCGTGGCGGCTGAAGCTGAAGCGCGCATGG 3674

QY 722 AGGCGCCCAAGTGAAGCAGTGGCGGCTGACCGAGGAGATCAAGGCGCTTGAACCGCA 781

Db 3675 ACGGCGCAAGTGAAGCAGTGGCGGCTGACCGAGGAGATCAAGGCGCTTGAAGGAGA 3734

QY 782 TCTGCGAGGATGAGAGGAGGCGCAAGATCACAAGATCGGCGCGCGAGACCGCTTACA 841

Db 3735 TCTGCGAGGATGAGAGGAGGCGCAAGATCACAAGATCGGCGCGCGAGACCGCTTACA 3794

QY 842 ACACCGCGCTTCCCGCATCAAGAGAGGAGCAGCAGCAGTGGCGGCTGGTGGACT 901

Db 3795 ACACCGCGCTTCCCGCATCAAGAGAGGAGCAGCAGCAGTGGCGGCTGGTGGACT 3854

QY 902 TCGGCGAGTGAACAAAGCGCACCCAGGACTTCTGGAGGTGCGAGTGGCGCATCCCCACC 961

Db 3855 TCGGCGAGTGAACAAAGCGCACCCAGGACTTCTGGAGGTGCGAGTGGCGCATCCCCACC 3914

QY 962 CCGCGCGCTGAAGAGAGAGACCGTGAACCGTGGAGCGTGGCGGCGAGCGCTTACTTCA 1021

Db 3915 CCGCGCGCTGAAGAGAGAGACCGTGAACCGTGGAGCGTGGAGCGCGAGCGCTTACTTCA 3974

QY 1022 GCGTCCCGCTGGAGGAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGCATCAACA 1081

Db 3975 GCGTCCCGCTGGAGAGGAGACTTCCGCAAGTACACCGCTTACCATCCCGAGCATCAACA 4034

QY 1082 ACGAGACCGCGGCGATCGCTACCAAGTGTCTGCGCGCGGCTGGAAGGCGAGC 1141

Db 4035 ACGAGACCGCGGCGATCGCTACCAAGTGTCTGCGCGCGGCTGGAAGGCGAGC 4094

QY 1142 CCAGCATCTTCAGAGCAGCATGACCAAGATCTTGGAGCGCTTCCGCGCGCGAGCGCG 1201

Db 4095 CCGCATCTTCAGTGCAGCATGACCAAGATCTTGGAGCGCTTCCGCGCGCGAGCGCG 4154

QY 1202 AGATCGTGTATCTACCA-----GGCGCGCGCTGTAGCTGGCGAGCGCTTGGAGTGGCG 1255

Db 4155 ACATCGTGTATCTACCAAGTGTACATGGACCGCTGTAGCTGGCGAGCGCTTGGAGTGGCG 4214

QY 1256 AGCAGCGCGCAAGATCGAGGAGTGTGCGCAAGCACTGTCTGGCTGGGCTTCAACACC 1315

Db 4215 AGCAGCGCAAGATCGAGGAGTGTGCGCGAGCACTGTCTGGCTGGGCTTCAACACC 4274

Qy	1316	CCGACAGAAGACACCAAGAAGAGCCCCCTTCTCTGCCCAT-----CGAGCTGCACCCCG	1369
Db	4275	CCGACAGAAGACACCAAGAAGAGCCCCCTTCTCTGTGATGGGTACGAGCTGCACCCCG	4334
Qy	1370	ACAAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAGAGAGCTGCACCGTGAACGACA	1429
Db	4335	ACAAAGTGGACCGTGCAGCCCATCGTGTCTGCCGAGAAGAGAGCTGCACCGTGAACGACA	4394
Qy	1430	TCCAGAAGCTGGTGGGCAAGCTGAACTGGGGCAGCCAGATCTACCCGGGATCAAGGTGC	1489
Db	4395	TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCCACCCAGATCTACGCCGCGCATCAAGGTGC	4454
Qy	1490	GCAGCTGTGCAGCTGTCTCCGCGGCGCAAGGCCCTGACCGACATCGTGCCTCTGACCG	1549
Db	4455	GCAGCTGTGCAAGCTGTCTCCGCGGCGCACCAAGGCCCTGACCGAGTGGTGCCTCTGACCG	4514
Qy	1550	AGGAGCGCAGCTGGAGCTGGCGGAGAACCGGAGAGTCTTCGCGGAGCCCGTGCACGGCG	1609
Db	4515	AGGAGCGCAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGAAGGAGCCCGTGCACGGCG	4574
Qy	1610	TGTACTACGACCCCAAGAGACCTGTGTGGCGGAGATCCAGAAGCAGGGCCACGACCAAGT	1669
Db	4575	TGTACTACGACCCCAAGAGACCTGATCGCGGAGATCCAGAAGCAGGGCCAGGGCCAGT	4634
Qy	1670	GGACCTACGAGATCTACGAGGACCCCTTCAGAGACCTGAAGACCGCAAGTACGCCAAGA	1729
Db	4635	GGACCTACGAGATCTACGAGGACCCCTTCAGAGACCTGAAGACCGGACGATACGCCGCA	4694
Qy	1730	TGGGCACCGCCCAACACGACGCTGAAGCAGCTGACCGAGCCGCTGCAGAAGATCGCCA	1789
Db	4695	TGAAGGCGCCACACACGACGCTGAAGCAGCTGACCGAGGCGCTGCAGAAGATCGCCA	4754
Qy	1790	TGGAGGACATCTGTGATCTGGGGCAGAGACCCCAAGTTCCGGCTGCCATTCAGAAGGAGA	1849
Db	4755	CCGAGAGCATCTGTGATCTGGGGCAGAGACCCCAAGTTTCAAGCTGCCATTCAGAAGGAGA	4814
Qy	1850	CCTGGGAGACCTGGTGGACCGCACTACTTGCAGGCCACCACTGGATCCCGAGTGGGAGTTCG	1909
Db	4815	CCTGGGAGGCTTGTGTGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCG	4874
Qy	1910	TGAACACCCCCCTTGTGTAAGCTGTGTGTACAGCTGGAGAGAGAGCCCATCATTCGGCG	1969
Db	4875	TGAACACCCCCCTTGTGTAAGCTGTGTGTACAGCTGGAGAGAGAGCCCATCATTCGGCG	4934
Qy	1970	CCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCCGGCT	2029
Db	4935	CCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGGAGACCAAGTCTGGGCAAGGCCGGCT	4994
Qy	2030	ACGTGACCGACCGGGCCGGCAGAGATGTGTGAGCCTGACCGAGACCAACCAACAGAGA	2089
Db	4995	ACGTGACCGACCGGGCCGGCAGAGAGTGGTGCCTTACCGACCAACCAACCAAGAGA	5054
Qy	2090	CCGAGCTGCAGGCCATTCAGCTGGCCCTGCAGGACAGCGGAGGAGTGAACATCTGTA	2149
Db	5055	CCGAGCTGCAGGCCATTCACCTTGGCCCTGCAGGACAGCGGCTTGAGTGAACATCTGTA	5114
Qy	2150	CCGACGCCAGTACGCCCTTGGGCATCATCCAGGCCACCGCCGACAAAGAGCAGAGCGAGC	2209
Db	5115	CCGACGCCAGTACGCCCTTGGGCATCATCCAGGCCACCGCCGACAAAGAGCAGAGCGAGC	5174
Qy	2210	TGGTGAACGAGATCATCGACGAGCTGATCAGAAGGAGAAGGTGTACTGTAGCTGGGTGC	2269
Db	5175	TGGTGAACGAGATCATCGACGAGCTGATCAGAAGGAGAAGGTGTACTGTAGCTGGGTGC	5234
Qy	2270	CCGCCCAAGAGGCATCGCGGCAACGACGACATCGCAAGCTGGTGTGAGCAAGGGCATCC	2329
Db	5235	CCGCCCAAGAGGCATCGCGGCAACGACGAGTGGACGGCTGGTGTGAGCCCGGCATCC	5294
Qy	2330	GCAAGTGTCTTCTCTGGACGGCATCGATGGCGGCATCGTGTACTACAGTGA	2381
Db	5295	GCAAGTGTCTTCTCTGGACGGCATCGACAAAGGCCACCGAGGAGGACACAGAA	5346

RESULT 6	AX427931	9169 bp	DNA	linear	PAT 20-JUN-2002
LOCUS	AX427931	Sequence 169 from Patent WO0232943.			
DEFINITION	AX427931				
ACCESSION	AX427931				
VERSION	AX427931.1	GI:21538018			
KEYWORDS		synthetic construct.			
SOURCE		synthetic construct			
ORGANISM		artificial sequences.			
REFERENCE	1				
AUTHORS	Huang Y. and Nabel G.J.				
TITLE	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic Immunization				
JOURNAL	Patent: WO 0232943-A 169 25-APR-2002;				
FEATURES	GOVERNMENT OF THE UNITED STATES (US)				
Source	Location/Qualifiers				
	1..9169				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note="plasmid pVR1012x/s containing HIV genes"				
BASE COUNT	2227 a 2668 c 2505 g 1769 t				
ORIGIN					
	Query Match	80.5%;	Score 1978.4;	DB 6;	Length 9169;
	Best Local Similarity	91.1%;	Pred. No. 2.3e-198;		
	Matches 2178;	Conservative 0;	Mismatches 186;	Indels 28;	Gaps 6;
Qy	14	TGGCGGAGGCGCATGAGCCAGGCGCACACGAGCGCCAACATCTCTGATGAGCGCGCAACTTCA	73		
Db	2967	TGGCGGAGGCGCATGAGCCAGGCTGCAGCAGCACCAACATCATGATGATGAGCGCGCAACTTCA	3026		
Qy	74	AGGGCCCCAAGGCCATCATCAAGTCTTCAACTGGGGCAAGGAGGGCCACATCGCCCGCA	133		
Db	3027	AGGGCCCAAGGCGCAT---CAAGTCTTCAACTTGGGGCAAGGAGGGCCACCTGGGCCCGCA	3083		
Qy	134	ACTGCGGCGGCCCCCGCCGCAAGAAGGCGTGTGGAAGTGGCGGAAGGAGGGCCACCAAGATGA	193		
Db	3084	ACTGCGGCGCCCTGCGCAAGAAGGCGTGTGGAAGTGGCGGAAGGAGGGCCACCAAGATGA	3143		
Qy	194	AGGACTGCACCGAGCGCCAGGCCAACTTCTTCGCGAGAGACTGGCCCTTCCCGCAGGGCA	253		
Db	3144	AGGACTGCACCGAGCGCCAGCGCTAA-TTTTTTAGGGAAGATGTGGCCCTTCCCGCAGGGCA	3202		
Qy	254	AGGCCGCGGAGTTGCCACGAGCAGAGAACCGGCCAACAGCCCCACACGAGCCGAGCTGC	313		
Db	3203	AGGCCAGGGAATTTTCTCAGAGCAGACCAAGAGCCACAGCCCCACAGAGAGAGACTTC	3262		
Qy	314	AGGTGCGGGG-----CGACAACCCCGAGGAGCGCGCGCCGAGCGCCAGCGCA---	364		
Db	3263	AGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAGAGCAGGAGCCGATAGACAAGGAACGTG	3322		
Qy	365	---CCCTGAACCTTCCCGCAGATCACCTCTGTGACGCGCCCTCGTGTGAGCATCAAGGTGG	421		
Db	3323	TATCCTTTAGCTTCCCTCAGATCATCTTTTGGCAGCGACCCCTCGTCACATTAAGATAG	3382		
Qy	422	GCGGCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGAGACACCCTGCTGGAGGAGA	481		
Db	3383	GCGGCAGCTGAAGGAGGCCCTTCTAGACACCGCGCGCGAGCACACCCTGCTGGAGGAGA	3442		
Qy	482	TGAGCGCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGCGGCATCGCGGCTTCATCAAGG	541		
Db	3443	TGAACCTTGCCTGGCGGCGCTGTGAAGCCCAAGATGATCGCGGCATCGCGGCTTCATCAAGG	3502		
Qy	542	TGCGCCAGTACGACACAGATCCTTGATCGAGATCTCGGGCAAGAAGGCCCATCGCACCCGTGC	601		
Db	3503	TGGGCCAGTACGACACAGATCCTTGATCGAGATCTGCGGCCCAAGGCCCATCGCACCCGTGC	3562		
Qy	602	TGATCGGGCCCCACCCCGGTGAACATCATCGCGCCGCAAGATGCTGACCCAGCTGGGCTGCA	661		
Db	3563	TGGTGGGGCCCCACCCCGGTGAACATCATCGCGCCGCAAGATGCTGACCCAGATCGGCTGCA	3622		


```
QY 662 CCCTGAACCTTCCCATCATGCCCATCGAGACCGTCCCTGAAGCTGAAGCCCGCATGG 721
Db 3623 CCCTGAACCTTCCCATCATGCCCATCGAGACCGTCCCTGAAGCTGAAGCCCGCATGG 3682
QY 722 ACGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGAGAGAGATCAAGGCCCTGACCGCA 781
Db 3683 ACGGCCCCAAGTGAAGCAGTGGCCCTGACCGAGAGAGAGATCAAGGCCCTGAGGAGA 3742
QY 782 TCTGGAGAGATGAGAGAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCTACA 841
Db 3743 TCTGCACGAGATGAGAGAGGAGGCAAGATCAGCAAGATCGGCCCGCGAGAACCCCTACA 3802
QY 842 ACACCCCGTGTGCCATCAAGAAAGAGAGACAGCACCAAGTGGCGCAAGCTGGGACT 901
Db 3803 ACACCCCGTGTGCCATCAAGAAAGAGAGACAGCACCAAGTGGCGCAAGCTGGGACT 3862
QY 902 TCCGCGAGCTGAACAAGCGCACCGAGGACTTCTGGAGGTGAGTGGGCTGAGTGGGACT 961
Db 3863 TCCGCGAGCTGAACAAGCGCACCGAGGACTTCTGGAGGTGAGTGGGCTGAGTGGGACT 3922
QY 962 CGCGCGGCTTGAAGAAGAGAGAGCGTGTGCTGTGGAGTGGCGACGCTACTTCA 1021
Db 3923 CGCGCGGCTTGAAGAAGAGAGAGCGTGTGCTGTGGAGTGGCGACGCTACTTCA 3982
QY 1022 GCGTGGCCCTGGACGAGGACTTCCGCAGTACACCGCTTCAACATCCCGAGCATCAACA 1081
Db 3983 GCGTGGCCCTGGACGAGGACTTCCGCAGTACACCGCTTCAACATCCCGAGCATCAACA 4042
QY 1082 ACGAGACCCCGGATCCGCTACCAAGTACAACGCTGTCCCGCCAGGCTGGAGGCGACCC 1141
Db 4043 ACGAGACCCCGGATCCGCTACCAAGTACAACGCTGTCCCGCCAGGCTGGAGGCGACCC 4102
QY 1142 CCAGCATCTCCAGAGAGCATGATACCAAGATTCCTGGAGCCCTTCCCGCCCGCAACCCCG 1201
Db 4103 CCAGCATCTCCAGAGAGCATGATACCAAGATTCCTGGAGCCCTTCCCGCAAGACCCCG 4162
QY 1202 AGATCGTGTATCAACA-----GCGCCCTGTAGTGGGAGGACCTGGAGATCGGCC 1255
Db 4163 ACATCGTGTATCAACAAGGACTTACGACCACTGTAGTGGGAGGACCTGGAGATCGGCC 4222
QY 1256 AGCACCGCGCAAGATCGAGGAGTGGCGAAGACCTGCTCGTGGGCTTACACACCC 1315
Db 4223 AGCACCGCAACATCGAGGAGTGGCGAAGACCTGCTCGTGGGCTTACACACCC 4282
QY 1316 CCGACAGAGACCCAGAGAGGAGCCCTTCTGCCCCAT-----CGAGCTGCACCCCG 1369
Db 4283 CCGACAGAGACCCAGAGAGGAGCCCTTCTGTTGGATGGGCTACGAGCTGCACCCCG 4342
QY 1370 ACAAAGTGAAGTGCAGCGCATCGAGCTGCCGAGAGAGGAGAGTGGACCGTGAACGACA 1429
Db 4343 ACAAAGTGAAGTGCAGCGCATCGTGTGCCGAGAGAGGAGAGTGGACCGTGAACGACA 4402
QY 1430 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 1489
Db 4403 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 4462
QY 1490 GCAGCTGTGAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGACATCTGCCCCCTGACCG 1549
Db 4463 GCAGCTGTGAAGCTGCTGGCGGCGCCAAAGGCCCTGACCGAGTGTGTCGCCCTGACCG 4522
QY 1550 AGGAGGCGGAGCTGGAGCTGGCGAGAACCGGAGATCTTCCGCGGAGCCGTCGACGGCG 1609
Db 4523 AGGAGGCGGAGCTGGAGCTGGCGAGAACCGGAGATCTTCCGAGGAGCCGTCGACGGCG 4582
QY 1610 TGTACTACGACCCAGCAAGGAGCTGTGGCGCGAGATCCCAAGAGCAGGCGCACGACAGT 1669
Db 4583 TGTACTACGACCCAGCAAGGAGCTGTATCCGCGAGATCCAGAGAGCAGGCGCACGCGAGT 4642
QY 1670 GGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCAAGA 1729
Db 4643 GGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCGCA 4702
QY 1730 TGGCGACCGCCACACCAAGCAGCTGAAGCAGCTGACCGAGGCGCTGCAGAGATCGCCA 1789
```

```
Db 4703 TGAAGGCGCCACACACACACGAGCTGAACAGCTGACCGAGCGCTGCAGAGATCGCCA 4762
QY 1790 TGGAGAGATCTGTATCTTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAAGAGA 1849
Db 4763 CCGAGAGATCTGTATCTTGGGCAAGACCCCAAGTTCAGCTGCCCATCCAGAAGAGA 4822
QY 1850 CCGTGGAGACCTGTGGACCGGACTACTGGCAGGCCACCTTGGATCCCCGAGTGGGAGTGC 1909
Db 4823 CCGTGGAGACCTGTGGACCGGAGTACTGGCAGGCCACCTTGGATCCCCGAGTGGGAGTGC 4882
QY 1910 TGAACACCCCCCTGTGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCGCGG 1969
Db 4883 TGAACACCCCCCTGTGTGAAGCTGTGTACAGCTGGAGAGAGAGCCCATCATCGCGG 4942
QY 1970 CCGAGACCTTACCTGTGTGGACCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCT 2029
Db 4943 CCGAGACCTTACCTGTGTGGACCGCGCCCAACCGGAGACCAAGATCGGCAAGCGCGCT 5002
QY 2030 ACGTCAACCGACCGGCGCGCAGAGATCGTGAAGCTTGAGCTGACCGAGACCAACCAAGA 2089
Db 5003 ACGTCAACCGACCGGCGCGCAGAGATCGTGAAGCTTGAGCTGACCGAGACCAACCAAGA 5062
QY 2090 CCGAGCTCAGGCGCATCCAGCTGGCCCTCCAGACAGCGCGCAGAGGTGAACATCGTGA 2149
Db 5063 CCGAGCTCAGGCGCATCCAGCTGGCCCTCCAGACAGCGCGCCTGGAGGTGAACATCGTGA 5122
QY 2150 CCGAGACCGAGTACGCCCTGGGCATCATCCAGGCCCGAGCCCGACAGAGCGAGCGAGC 2209
Db 5123 CCGAGACCGAGTACGCCCTGGGCATCATCCAGGCCCGAGCCCGACAGAGCGAGCGAGC 5182
QY 2210 TGGTGAACCAAGATCGAGCAGCTGATCAAGAGAGAGAGTGTACCTGAGCTGGGTGC 2269
Db 5183 TGGTGAACCAAGATCGAGCAGCTGATCAAGAGAGAGAGTGTACCTGAGCTGGGTGC 5242
QY 2270 CCGGCCCAAGAGGATCGCGCGCACAGCAGATGACAAAGCTGTGAGAGAGGCGATCC 2329
Db 5243 CCGGCCCAAGAGGATCGCGCGCACAGCAGAGTGTGAGCGGCTGTGAGCGCGCGATCC 5302
QY 2330 GCAAGGTGCTTCTCTGGACCGCATCGATGGCGGATCGTGTATCTACAGTA 2381
Db 5303 GCAAGGTGCTTCTCTGGACCGCATCGATGGCGGATCGATGGCGGAGGAGCAGAGAA 5354
```

```
RESULT 7
AX427926 LOCUS AX427926 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 164 from Patent WO0232943.
ACCESSION AX427926
VERSION AX427926.1 GI:21538013
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
JOURNAL genetic immunization
PATENT: WO 0232943-A 164 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
LOCATION/Qualifiers
source 1. .9194
/organism="synthetic construct"
/dbxref="taxon:32630"
/notes="plasmid pVR1012x/s containing HIV genes"
```

```
BASE COUNT 2232 a 2672 c 2523 g 1767 t
ORIGIN
```

```
Query Match 80.4%; Score 1975; DB 6; Length 9194;
Best Local Similarity 90.9%; Pred. No. 5.2e-198;
Matches 2177; Conservative 0; Mismatches 190; Indels 28; Gaps 6;
QY 14 TGGCGAGGCGCATGAGCCAGG---CCACCAGCGCAACATCCTGATGCAGCGCAGCAACT 70
```


Db 2967 TGGCCGAGCCATGAGCGAGTACCAACAGCGCCACCATCATGATGCGAGCGGCAACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAATGCTTCAACTTCGCGCAAGGAGGGCCACATCGCCC 130
Db 3027 TCCGCAACAGCGCAAGTCTGAAGTCTCAACTTCGCGCAAGGAGGGCCACACCGCCC 3086
QY 131 GCAACTTGGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGGAGGGCCACACAGA 190
Db 3087 GCAACTTGGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCCAGCCCAACTTCTTCGCGCAGGAGACCTTGGCCCTTCCCCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGACAGGCTAA -TTTTTATGGAAGATCTGGCCCTTCCCCAAG 3205
QY 251 GCAAGGCCCGGAGTTCGCCAGCGAGCAGAGAACCGCGGCCACACGCCCCACCGCGGAGC 310
Db 3206 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACAGAGCCCAACAGCCCCACACAGAGAGC 3265
QY 311 TGCAGGTGCGCGG - - - - -CGACAACCCCGCAGCGAGGCGCGCGCGAGCGCAGGGCA 364
Db 3266 TTCAAGTTTGGGAAGAGAGACAACACTCCCTCTCAGAGCAGAGCGCGATAGACAGGAA 3325
QY 365 - - - - -CCCTGAATTCCTCCCGAGATCACCTGTGGCAGCGCCCTTGGTGAAGCATCAAGG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCTGTCAATAAAGA 3385
QY 419 TGGCGCCGAGATCAAGGAGGCCCTGCTGCACACCGCGCGCGAGCACACGCTGTGAGG 478
Db 3386 TAGGGGCGAGCTGAAGGAGGGCCCTTCTAGACACCGCGCGCGAGCACACGCTGTGAGG 3445
QY 479 AGATGAGCCTGCCCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGGCTTCAATCA 538
Db 3446 AGATGAACTTGCCCGCGCTGGAAGCCCAAGATGATCGCGGGCATCGCGGGCTTCAATCA 3505
QY 539 AGGTGCGCCAGTAGACCAAGATCCTGTATCGAGATCTTCGGGCAAGAGCGCATGCGCACCG 598
Db 3506 AGGTGCGCCAGTAGACCAAGATCCTGTATCGAGATCTTCGGGCAAGAGCGCATGCGCACCG 3565
QY 599 TGTGATGCGCCACCCCGTGAACATCATCGSCGCCAATCATGCTGACCCAGCTGGGCT 658
Db 3566 TGCTTGGTGGCCCCCAGCGTGAACATCATCGSCGCCAATCATGCTGACCCAGATCGGCT 3625
QY 659 GCACCTCTGAACTTCCCATCAGCCCATCGAGACCGCTGCCGTTGAAGCTGAAGCCCGGCA 718
Db 3626 GCACCTCTGAACTTCCCATCAGCCCATCGAGACCGCTGCCGTTGAAGCTGAAGCCCGGCA 3685
QY 719 TGGAGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGAGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTTGGTGG 3745
QY 779 CCATCTCGAGGAGATGAGAGAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCT 838
Db 3746 AGATCTGACCGAGATGAGAGAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCT 3805
QY 839 ACAACACCCCGTGTTCGCCATCAAGAGAGAGAGACACCAAGTGGCGCAAGCTGTGG 898
Db 3806 ACAACACCCCGTGTTCGCCATCAAGAGAGAGAGACACCAAGTGGCGCAAGCTGTGG 3865
QY 899 ACTTCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC 958
Db 3866 ACTTCGGGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGTGGGCATCCCC 3925
QY 959 ACCCGCGGCTGAAGAAGAAGAGAGCTGACCGTGTGACGTGGGCGACGCCCTACT 1018
Db 3926 ACCCGCGGCTGAAGAAGAAGAGAGCTGACCGTGTGACGTGGGCGACGCCCTACT 3985
QY 1019 TCAGCGTCCCTTGAGCAGAGACTTTCGCAAGTACACCGCCCTTCAACATCCCCAGCATCA 1078
Db 3986 TCAGCGTCCCTTGAGCAGAGACTTTCGCAAGTACACCGCCCTTCAACATCCCCAGCATCA 4045
QY 1079 ACAAGGAGACCCCGGATCCGCTACCAAGTACAGCTGCTGCCCGCAGGCTGGAAGGCA 1138

Db 4046 ACAAGAGACCCCGCATCCGCTACCAAGTACAACTGTGCTCCCGAGGGCTGGAAGGGCA 4105
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCGCCGCGCATCTTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CCAGAGTCTGTATCTACCA - - - - -GGCCCCCTGTACGTGGGCGAGGACCTGGAGATCG 1252
Db 4166 CCAGATCTGTATCTACCAAGTACATGACCACTGTACGTGGGCGAGGACCTGGAGATCG 4225
QY 1253 GCAGACACCGCGCAGATGAGAGCTGCGCAAGCCTGCTGCGCTGGGGCTTACCA 1312
Db 4226 GCAGACACCGCGCAGATGAGAGCTGCGCAAGCCTGCTGCGCTGGGGCTTACCA 4285
QY 1313 CCGCCGACAAGAGCAGCAGAAAGGAGCCCTTCTCTGCCCAT - - - - -GGAGTGCACC 1366
Db 4286 CCGCCGACAAGAGCAGCAGAAAGGAGCCCTTCTCTGTGATGGGCTACGAGTGCACC 4345
QY 1367 CCGACAAGTGGACCGTGCAGCCCATGAGCTGCGCGAGAGAGAGAGCTGAGACCTGAACG 1426
Db 4346 CCGACAAGTGGACCGTGCAGCCCATGAGCTGCGCGAGAGAGAGAGCTGAGACCTGAACG 4405
QY 1427 ACATCCAGAGACTGGTGGGCAAGCTGAACCTGGCGCAGCAGATCTACCCGGCATCAAGG 1486
Db 4406 ACATCCAGAGACTGGTGGGCAAGCTGAACCTGGCGCAGCAGATCTACCCGGCATCAAGG 4465
QY 1487 TGGCCGAGCTGTGCAAGCTGCTGCGCGCGCCAAAGGCCCTGACCGAGATCTGTCGCCCTGA 1546
Db 4466 TGGCCGAGCTGTGCAAGCTGCTGCGCGCGCCAAAGGCCCTGACCGAGGTGTCGCCCTGA 4525
QY 1547 CCGAGGAGCGGAGCTGGAGCTGGCGCGAGAACCCCGAGATCTCTGCGGAGGCCCTGTGACG 1606
Db 4526 CCGAGGAGCGGAGCTGGAGCTGGCGCGAGAACCCCGAGATCTCTGAAAGGAGGCCCTGTGACG 4585
QY 1607 GCGTGTACTAGCACCCCGAGCAAGGACCTGTGTGCGCGAGATCCAGAAGCAGGCGCCAGCAC 1666
Db 4586 GCGTGTACTAGCACCCCGAGCAAGGACCTGTGTGCGCGAGATCCAGAAGCAGGCGCCAGGCGC 4645
QY 1667 AGTGGAGCTACAGATCTACCGAGGCGCTTCAAGAACCCTGAAGACCGGCAAGTACGCCA 1726
Db 4646 AGTGGAGCTACAGATCTACCGAGGCGCTTCAAGAACCCTGAAGACCGGCAAGTACGCCA 4705
QY 1727 AGATGCGCAGCGCCACACCAAGCAGCTGAGCAGCTGACCGAGGCGCTGCGAGAGATCG 1786
Db 4706 GCATGAGAGGGCGCCACCAACAGCAGCTGAAGCAGCTGACCGAGGCGCTGCGAGAGATCG 4765
QY 1787 CCATGGAGAGCATCTGATCTGGGCGAAGACCCCAAGTTCGGCTGCCATCCAGAAAG 1846
Db 4766 CCACCGAGAGCATCTGATCTGGGCGAAGACCCCAAGTTCAGCTGCCATCCAGAAAG 4825
QY 1847 AGACTGGGAGACTGTGTGGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGGAGT 1906
Db 4826 AGACTGGGAGCGCTGTGTGGACCGAGTACTGGCAGGCCACCTGGATCCCGAGTGGGAGT 4885
QY 1907 TCGTGAACACCCCGCTTGTGTGAAGCTGTGTACCAAGCTGAGAGAGAGCCCATCATCG 1966
Db 4886 TCGTGAACACCCCGCTTGTGTGAAGCTGTGTACCAAGCTGAGAGAGAGCCCATCATCG 4945
QY 1967 GCGCGGAGACCTTCTACGTGGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGG 2026
Db 4946 GCGCGGAGACCTTCTACGTGGACGCGCGCCCAACCGGAGACCAAGATCGGCAAGGCGG 5005
QY 2027 GCTACGTGACCGACCGGGCGCGGAGAGATCTGTGAGCTTACCGAGACCAACACAGAG 2086
Db 5006 GCTACGTGACCGACCGCGCGCGGAGAGTGTGTGCTGAGAGAGAGCCCATCATCG 5065
QY 2087 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCGAGGAGAGAGTGAACATCG 2146
Db 5066 AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCGAGGAGAGAGTGAACATCG 5125
QY 2147 TGACCGACGAGCAGTACCGCTTGGGCATCTACAGGCCCGCCAGCCAGAGAGAGAGG 2206
Db 5126 TGACCGACGAGCAGTACCGCTTGGGCATCTACAGGCCCGCCAGCCAGAGAGAGGCG 5185

QY 2207 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 2266
|||||
Db 5186 AGCTGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAGGTGTACCTGAGCTGGG 5245
|||||
QY 2267 TCCCGCGCCACAAGGCGATCGCGGCAAGGAGCAGATCGACAAGCTGGTGAGCAAGGCA 2326
|||||
Db 5246 TCCCGCGCCACAAGGCGATCGCGGCAAGGAGCAGATCGACAAGCTGGTGAGCGCGGCA 5305
|||||
QY 2327 TCCGCAAGGTGCTGTCTCTGACGCGATCGATGGCGGATCGTGATCTACAGTA 2381
|||||
Db 5306 TCCGCAAGGTGCTGTCTCTGACGCGATCGACAAGGCCAGGAGGAGCAGGAA 5360
|||||
RESULT 8
AX427925
LOCUS AX427925 9194 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 163 from Patent WO0232943.
ACCESSION AX427925
VERSION AX427925.1 GI:21538012
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
1 artificial sequences.
REFERENCE
AUTHORS Huang,Y. and Nabel,G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 163 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1..9194
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2232 a 2671 c 2524 g 1767 t
ORIGIN
Query Match 80.3%; Score 1973.4; DB 6; Length 9194;
Best Local Similarity 90.9%; Pred. No. 7.6e-198;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;
QY 14 TGGCGGAGCCATGAGCCAGG---CCAGCAGCGCCACAGATCCTGATGAGCGCGCAACT 70
|||||
Db 2967 TGGCGGAGCCATGAGCGAGGTGACCAACAGCGCCACCATGATGAGCGCGCAACT 3026
|||||
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGGCGCAAGGAGGCGCACATCGCCC 130
|||||
Db 3027 TCCGCAACCGCGCAAGATCTGAAGTGTCTCAACTGGCGCAAGGAGGCGCACACCGCCC 3086
|||||
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGCAAGGAGGCGCCACACAGA 190
|||||
Db 3087 GCAACTGCCCGCGCCCGCCGCAAGAGGCGTCTGGAAGTGGCGCAAGGAGGCGCCACACAGA 3146
|||||
QY 191 TGAAGGACTGACCGAGGCGCCAGCGCAACTTCTTCCGCGAGGAGCTGGCCCTTCCCGCAGG 250
|||||
Db 3147 TGAAGGACTGACCGAGGCGCAAGCGCTAA-TTTTATTAGGAGATCTGGCCCTTCCACAAAG 3205
|||||
QY 251 GCAAGGCGCGGAGTTCCCGCAGCGAGCAGAACCGCGCCACAGCCCGCCAGCGCGAGC 310
|||||
Db 3206 GGAAGGCCAGGGAATTTCTTCAGAGCAGACCCAGAGCCACAGCCCGCCACAGAGAGAGC 3265
|||||
QY 311 TGCAGGTGCGCGG-----CGACAACCGCCGACGCGAGGCGCGCGCGAGGCGCGAGGCA 364
|||||
Db 3266 TTCAGGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAGGAGGAGCGCATAGACAAGAA 3325
|||||
QY 365 -----CCCTGAACCTTCCCGAGATCACTGTGTGGCAGCGCCCGCTGGTGAGCATAAGG 418
|||||
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACATAAAGA 3385
|||||
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGTGGACACCGCGCGCGAGCACACCGTGTGGAGG 478
|||||
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCTTTAGACACCGCGCGCGAGCACACCGTGTGGAGG 3445
|||||

QY 479 AGATGAGCTGTCCCGCAGTGTGAAGCCCAAGATGATCGCGGCATCGCGCGCTTCATCA 538
|||||
Db 3446 AGATGAACCTGCCCAGCGCTGGAAGCCCAAGATGATCGCGGCATCGCGCGCTTCATCA 3505
|||||
QY 539 AGGTGCGCCAGTACGACCAAGATCTCTGATGAGATCTCGCGCAAGAGGCGCATCGGCACCG 598
|||||
Db 3506 AGGTGCGCCAGTACGACCAAGATCTCTGATGAGATCTCGCGCCACAGGCGCATCGGCACCG 3565
|||||
QY 599 TGCTGATCGGCGCCACCGCTGAGATCATCGCGCGCAACATGATGACCCAGCTGGGCT 658
|||||
Db 3566 TGCTGATCGGCGCCACCGCTGAGATCATCGCGCGCAACCTGCTGACCCAGCTGGCT 3625
|||||
QY 659 GCACCTGAACTTCCCGCATCGCGCCATCGAGACGCTGCCCGTGAAGCTGAAGCCCGGCA 718
|||||
Db 3626 GCACCTGAACTTCCCGCATCGCGCCATCGAGACGCTGCCCGTGAAGCTGAAGCCCGGCA 3685
|||||
QY 719 TGGACGCGCCCAAGGTGAAGTGAAGTGGCGCTGACCGAGGAGAGATCAAGAGCCCTGACCG 778
|||||
Db 3686 TGGACGCGCCCAAGGTGAAGTGAAGTGGCGCTGACCGAGGAGAGATCAAGAGCCCTGGTGG 3745
|||||
QY 779 CCATCTGCGAGGAGATGGAGAGGAGGCGCAAGATCACCAAGATCGCGCCCGAGAACCCCT 838
|||||
Db 3746 AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAGCAAGATCGCGCCCGAGAACCCCT 3805
|||||
QY 839 ACAACACCGCCCGTGTTCGCCATCAAGAAGAGGAGCAGCACCAAGTGGCGCAAGCTGGTGG 898
|||||
Db 3806 ACAACACCGCCCGTGTTCGCCATCAAGAAGAGGAGCAGCACCAAGTGGCGCAAGCTGGTGG 3865
|||||
QY 899 ACTTCCGCGAGGTGAACAAGCGCACCCAGGACTTCTGGAGGTGAGCTGGCGCATCCCC 958
|||||
Db 3866 ACTTCCGCGAGGTGAACAAGCGCACCCAGGACTTCTGGAGGTGAGCTGGCGCATCCCC 3925
|||||
QY 959 ACCCGCGCGCTGAAAGAGAGAGAGCGTGACCGTGTGACGCTGGCGAGCGCTACT 1018
|||||
Db 3926 ACCCGCGCGCTGAAAGAGAGAGAGCGTGACCGTGTGACGCTGGCGAGCGCTACT 3985
|||||
QY 1019 TCAGGTGCCCTTGGACGAGGACTTCCCGCAAGTACACCGCTTCCACCTCCCGCAGCATCA 1078
|||||
Db 3986 TCAGGTGCCCTTGGACGAGGACTTCCCGCAAGTACACCGCTTCCACCTCCCGCAGCATCA 4045
|||||
QY 1079 ACAAGGAGACCCCGGATCCGCTACCAAGTACAAGTGTGCCCCAGGCGCTGGAAGGCA 1138
|||||
Db 4046 ACAAGGAGACCCCGGATCCGCTACCAAGTACAAGTGTGCCCCAGGCGCTGGAAGGCA 4105
|||||
QY 1139 GCGCCAGATCTTCCAGAGCAGCATGACCAAGATCCTTGAGCCCTTCCGCGCCCGCAAGC 1198
|||||
Db 4106 GCGCCAGATCTTCCAGTGCAGCATGACCAAGATCCTTGAGCCCTTCCGCGCCCGCAAGC 4165
|||||
QY 1199 CCGAGATCGTGATCTACCA-----GGCGCCCTGAGTGGCGAGCGACCTGGAGATCG 1252
|||||
Db 4166 CCGAGATCGTGATCTACCAAGTACATGAGCGACCTGAGTGGCGAGCGCTGGAGATCG 4225
|||||
QY 1253 GCCAGACCGCGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGCTGGGCTTCACCA 1312
|||||
Db 4226 GCCAGACCGCGCCCAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGCTGGGCTTCACCA 4285
|||||
QY 1313 CCGCCGACAAAGACCCAGAGAGGCGCCCTTCTTCCGCGCAT-----CGAGCTGCAC 1366
|||||
Db 4286 CCGCCGACAAAGACCCAGAGAGGCGCCCTTCTTCCGCGCAT-----CGAGCTGCAC 4345
|||||
QY 1367 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCCGAGAGGAGCTGGACCGTGAAGC 1426
|||||
Db 4346 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCCGAGAGGAGCTGGACCGTGAAGC 4405
|||||
QY 1427 ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGCGCAGCAGATCTACCCCGCATCAAGG 1486
|||||
Db 4406 ACATCCAGAAGCTGGTGGGCAAGCTGAACTGGCGCAGCAGATCTACCCCGCATCAAGG 4465
|||||
QY 1487 TGCGCGCAGCTGTGCAAGCTGTGCGCGCGCCAGGCGCTTGACCGACATCTGTGCCCTTGA 1546
|||||
Db 4466 TGCGCGCAGCTGTGCAAGCTGTGCGCGCGCCAGGCGCTTGACCGAGCTGTGCCCTTGA 4525
|||||

QY	1547	CCGAGGAGCGGAGCTGGAGCTGGCCGAGAAACCGGAGATCTCTCGCGAGCCCGCTGCACG	1606
Db	4526	CCGAGGAGCGGAGCTGGAGCTGGCCGAGAAACCGGAGATCTCTGAAGGAGCCCGCTGCACG	4585
QY	1607	GCGTGACTACGACCCAGCAAGGACCTGGTGCCGAGATCCAGAGACGAGGCGACGACC	1666
Db	4586	GCGTGACTACGACCCAGCAAGGACCTGTATCGCCGAGATCCAGAGACGAGGCGCAGGGCC	4645
QY	1667	AGTGGACCTACGAGATCTACGAGGAGCCCTTCAAGAACCTGGAAGACCGGCAAGTACGCCA	1726
Db	4646	AGTGGACCTACGAGATCTACGAGGAGCCCTTCAAGAACCTGGAAGACCGGCAAGTACGCCC	4705
QY	1727	AGATGCGACCGCCACACAGCAGTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG	1786
Db	4706	GCATGAAGGCGCCACACAGCAGTGAAGCAGCTGACCGAGGCGCTGCAGAAAGATCG	4765
QY	1787	CCATGAGAGCATCTGTATCTGGGGCAAGACCCCAAGTTCCGCGCTGCCCATCCAGAAGG	1846
Db	4766	CCACCGAGAGCATCTGTATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAGG	4825
QY	1847	AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCGCCACTGGATCCCGAGTGGGAGT	1906
Db	4826	AGACCTGGGAGGCTGGTGACCGAGTACTTGGCAGGCGCCACTGGATCCCGAGTGGGAGT	4885
QY	1907	TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGAGCCCATCATCG	1966
Db	4886	TCGTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGAGCCCATCATCG	4945
QY	1967	GCGCGAGACCTTCTAGCTGGAGCGCGCCGCAACCGCGAGACCAAGATCGGCAAGGCCG	2026
Db	4946	GCGCGAGACCTTCTAGCTGGAGCGCGCGCCGCAACCGCGAGACCAAGCTGGGCAAGGCCG	5005
QY	2027	GCTACGTGACCGAGCGGCGCGCAGAGATCTGTAGCCCTGACCGAGACCAACCAACGAGA	2086
Db	5006	GCTACGTGACCGAGCGGCGCGCAGAGATCTGTAGCCCTGACCGAGACCAACCAACGAGA	5065
QY	2087	AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCGAGGACAGCGGAGGAGTGAACATCG	2146
Db	5066	AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCGAGGACAGCGGAGTGAACATCG	5125
QY	2147	TGACCGACACCGAGTACGCCCTGGGCATCATCCAGGCGCCAGCCGACCAAGAGAGCG	2206
Db	5126	TGACCGACACCGAGTACGCCCTGGGCATCATCCAGGCGCCAGCCGACCAAGAGAGAGCG	5185
QY	2207	AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG	2266
Db	5186	AGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACCTGAGCTGGG	5245
QY	2267	TGCCCCCGCCACAGGGGATCGGGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGGCA	2326
Db	5246	TGCCCCCGCCACAGGGGATCGGGGCAACGAGCAGTGGAGCTGGTGAGCGCCGCGCA	5305
QY	2327	TCCGCAAGGTGCTTCTCTGGACGGCATCGATGCGCGCATCGTGTATCCAGTA	2381
Db	5306	TCCGCAAGGTGCTTCTCTGGACGGCATCGACAAAGGCCCGAGGAGGACGACGAGAA	5360
RESULT 9			
AX427927			
LOCUS	AX427927	12411 bp	DNA
DEFINITION	Sequence 165 from Patent WO0232943.	linear	PAT 20-JUN-2002
ACCESSION	AX427927		
VERSION	AX427927.1	GI:21538014	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	artificial sequences.		
AUTHORS	1		
TITLE	Huang, Y. and Nabel, G.J.		
JOURNAL	Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization		
	Patent: WO 0232943-A 165 25-APR-2002;		
	GOVERNMENT OF THE UNITED STATES (US)		

FEATURES	Location/Qualifiers	
source	1. .12411	
	/organism="synthetic construct"	
	/db_xref="taxon:32630"	
	/note="plasmid pVR1012x/s containing HIV genes"	
BASE COUNT	2938 a 3733 c 3476 g 2264 t	
ORIGIN		
	Query Match 80.3%; Score 1973.4; DB 6; Length 12411;	
	Best Local Similarity 90.9%; Pred. No. 7.1e-198;	
	Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;	
QY	14	TGCGCGAGGCGCATGAGCCAGG---CCACACAGCGCCAACTCTCTGATCAGCGCAGCAACT 70
Db	2967	TGCGCGAGGCGCATGAGCCAGGCGTACCAACAGCGCCACCATCATGATCAGCGCGGCAACT 3026
QY	71	TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTCGGCAAGGAGGGCCACATCGCC 130
Db	3027	TCGCGCAACCAAGCGCATGCTGAAGTGTCTCAACTCGGCAAGGAGGGCCACACCGCCC 3086
QY	131	GCAACTGCCCGCGCCCGCCCAAGAGGGCTGTGGAAAGTGCAGCAAGGAGGGCCACACAGA 190
Db	3087	GCAACTGCCCGCGCCCGCCCAAGAGGGCTGTGGAAAGTGCAGCAAGGAGGGCCACACAGA 3146
QY	191	TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCCGAGGACCTGGCCTTCCCCGAGG 250
Db	3147	TGAAGGACTGCACCGAGCGGACAGGCTAA-TTTTTTAGGGAAGATCTGGCCTTCCCAACAG 3205
QY	251	GCAAGGCCCGCGAGTTCCCGAGCGGAGAGAACCGCGCCCAACAGCCCGCCACCGCGCGAGC 310
Db	3206	GGAAGGCCAGGGAATTTTCTTCAGAGCAGGACCAAGAGCCCAACAGCCCGCCACCAAGAGAGC 3265
QY	311	TGACAGTGGCCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGAGCGCCAGGGCA 364
Db	3266	TTAGGTTTGGGAGAGAGACAACAACCTCCCTCTCAGAAAGCAGGAGCGGATAGACAAGAA 3325
QY	365	-----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTCGGTGAGCATCAAGG 418
Db	3326	CTGTATCTTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGTCACAATAAGA 3385
QY	419	TGGCGCGCGCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGAGCAGACCTGTCTGGAGG 478
Db	3386	TAGGGGCGCAGCTGAAGAGGCGCCCTCTAGACACCGCGCGCGAGCACCTGTCTGGAGG 3445
QY	479	AGATGAGCCTTCCCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 538
Db	3446	AGATGAACTTCCCGCGCGCTGGAAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 3505
QY	539	AGGTGCGCGCAGTAGGACAGATCCTGATCGAGATCTCGGGCAAGAGCCATCGGCACCG 598
Db	3506	AGGTGGGCGCAGTAGGACAGATCCTGATCGAGATCTCGGGCCACAAGGCCATCGGCACCG 3565
QY	599	TGCTGATCGCGCCCGCCCGTGAACATCATCGCGCCCAACATCTGACCCAGCTGGGCT 658
Db	3566	TGCTGGTGGGCGCCCGCCCGTGAACATCATCGCGCCCAACATCTGACCCAGATCGGCT 3625
QY	659	GCACCTTGAACCTTCCCGCATCAGCCCGCATCGAGACCGTGGCCGTGAAGTGAAGCCCGGCA 718
Db	3626	GCACCTTGAACCTTCCCGCATCAGCCCGCATCGAGACCGTGGCCGTGAAGTGAAGCCCGGCA 3685
QY	719	TGGAGCGCCCGCAAGGTGAAGCAGTGGCGCCCTGACCGAGGAGAAATCAAGGCCCTGACCG 778
Db	3686	TGGAGCGCCCGCAAGGTGAAGCAGTGGCGCCCTGACCGAGGAGAAATCAAGGCCCTTGTGG 3745
QY	779	CCATCTCGGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGCAAGCCCT 838
Db	3746	AGATCTGACCGGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGGCAAGCCCT 3805
QY	839	ACAAACCGCCCGTTCGCCCATCAAGAGAGGAGACAGCACCAAGTGGCGCAAGCTGTGG 898
Db	3806	ACAAACCGCCCGTTCGCCCATCAAGAGAGGAGACAGCACCAAGTGGCGCAAGCTGTGG 3865
QY	899	ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958

Db 3866 ACTTCGGGAGCTGAACAAGCGCACCCAGAGACTTCTGGGAGGTGACGTGGGCATCCCCC 3925
QY 959 ACCCCGCGGGCTGAAGAAGAAGAGAGCTGACCGTCTGCTGACGTGGGGCAGCCCTACT 1018
Db 3926 ACCCCGCGGGCTGACGACGAGAGAGAGCTGACCGTCTGCTGACGTGGGGCAGCCCTACT 3985
QY 1019 TCAGCGTGCCCTGGAGAGGACTTCCGCAAGTACACCGCCCTTACCACATCCCCAGCATCA 1078
Db 3986 TCAGCGTGCCCTGGACAAGACTTCCGCAAGTACACCGCCCTTACCACATCCCCAGCATCA 4045
QY 1079 ACAAGGAGACCCCGGATCCGCTACCAAGTACACAGTGTCTGCCAGGGCTGGAGGGCA 1138
Db 4046 ACAAGGAGACCCCGGATCCGCTACCAAGTACACAGTGTCTGCCAGGGCTGGAGGGCA 4105
QY 1139 GCCCGAGACTTTCAGAGAGCATGATCAAGATCCCTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4106 GCCCGGCACTTTCAGTGACATGATACCAAGATCCCTGGAGCCCTTCCGCAAGCAGAAC 4165
QY 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTGTAGTGGGAGGACCTGGAGATCG 1252
Db 4166 CCGACATCGTGATCTACCAAGTACATGACGACCTGTAGTGGGAGGACCTGGAGATCG 4225
QY 1253 GCCAGCACCGCGCAAGATCGAGAGTGTGCGAAGCACCCTGCTGGCTGGGGCTTCACCA 1312
Db 4226 GCCAGCACCGCGCAAGATCGAGAGTGTGCGAAGCACCCTGCTGGCTGGGGCTTCACCA 4285
QY 1313 CCCCAGCAAGAACACAGAGGAGCGCCCTTCTTCCCTGCCCCAT-----CGAGCTGCACC 1366
Db 4286 CCCCAGCAAGAACACAGAGGAGCGCCCTTCTTCCCTGCCCCAT-----CGAGCTGCACC 4345
QY 1367 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCGCCGAGAGAGAGCTGGACCTGAACG 1426
Db 4346 CCGACAAGTGACCGCTGCAGCCCATCGAGCTGCGCCGAGAGAGAGCTGGACCTGAACG 4405
QY 1427 ACATCCAGAGCTGGTGGGCAAGCTGACCTGGGCGCAGCAGATCTACCCCGCATCAAGG 1486
Db 4406 ACATCCAGAGCTGGTGGGCAAGCTGACCTGGGCGCAGCAGATCTACCCCGCATCAAGG 4465
QY 1487 TGCGCCAGCTGTCAAGCTGTGCGCGCGCAAGGCGCTGACCGACATCTGCCCCCTGA 1546
Db 4466 TGCGCCAGCTGTCAAGCTGTGCGCGCGCAAGGCGCTGACCGAGGTGTGCCCTGA 4525
QY 1547 CCGAGGAGCGGAGCTGGAGCTGCGCGAGAACCGCGAGATCTCTCGGAGCGCCGTGCACG 1606
Db 4526 CCGAGGAGCGGAGCTGGAGCTGCGCGAGAACCGCGAGATCTCTGAAGGAGCGCCGTGCACG 4585
QY 1607 GCGTGTACTAGCACCCAGCAAGACCTGTGTGGCGGAGATCCAGAACGAGGCGCACGACC 1666
Db 4586 GCGTGTACTAGCACCCAGCAAGACCTGTGTGGCGGAGATCCAGAACGAGGCGCACGACC 4645
QY 1667 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGGCA 1726
Db 4646 AGTGGACCTACAGATCTACAGGAGCGCTTCAAGAACCTGAAGACCGGCAAGTACGGCC 4705
QY 1727 AGATGCGACCGCCACACCAACGACGTGAAGAGCTGACGAGGCGGTGCGAAGATCG 1786
Db 4706 GCATGAAGGCGGCCACACCAACGACGTGAAGAGCTGACGAGGCGGTGCGAAGATCG 4765
QY 1787 CCATGGAGAGCATGCTGATCTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 1846
Db 4766 CCACGAGAGCATGCTGATCTGGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 4825
QY 1847 AGACTGGGAGACTGTGTGAGCCGACTACTTGGCAGGCGCACTTGGATCCCGAGTGGGAGT 1906
Db 4826 AGACTGGGAGCTGTGTGAGCCGAGTACTTGGCAGGCGCACTTGGATCCCGAGTGGGAGT 4885
QY 1907 TCGTGAACACCCCGCTGTGTGAAGTGTGTACCAAGCTGGAGAAGAGGCCCATCATCG 1966
Db 4886 TCGTGAACACCCCGCTGTGTGAAGTGTGTACCAAGCTGGAGAAGAGGCCCATCATCG 4945
QY 1967 GCGCGGAGACTTCTACGTGACGGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCCG 2026

Db 4946 GCGCGGAGACCTTCTACGTGAGCGCGCGCCCAACCGCGAGACCAAGCTGGGCAAGGCCG 5005
QY 2027 GCTAGCTGACCGACCGCGCGCGGAGAGAGTCTGAGCTGACCGAGACCAACCAACGAGA 2086
Db 5006 GCTAGCTGACCGACCGCGCGCGGAGAGAGTCTGAGCTGACCGAGACCAACCAACGAGA 5065
QY 2087 AGACCGAGCTGCAGGCGCATTCAGCTGGCCCTGCGAGGACAGCGGCGAGGTGAACATCG 2146
Db 5066 AGACCGAGCTGCAGGCGCATTCAGCTGGCCCTGCGAGGACAGCGGCGTGGAGGTGAACATCG 5125
QY 2147 TGACCGAGACCGAGTACGCGCTTGGGCATCATCCAGGCGCGAGCCCGACAGAGGAGCGG 2206
Db 5126 TGACCGAGACCGAGTACGCGCTTGGGCATCATCCAGGCGCGAGCCCGACAGAGGAGCGG 5185
QY 2207 AGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTGGG 2266
Db 5186 AGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAAGGAGAGTGTACCTGAGCTGGG 5245
QY 2267 TGCCCGCGCCCAAGGCGCATTCGCGGCAACGAGCAGATCGACAAGCTGTGTAGCAAGGCA 2326
Db 5246 TGCCCGCGCCCAAGGCGCATTCGCGGCAACGAGCAGCTGTGTAGCGCGCGCA 5305
QY 2327 TCCGCAAGGTGCTGTCTTGGACGCGCATCGATGGCGGCGATCGTGATCTACCAGTA 2381
Db 5306 TCCGCAAGGTGCTGTCTTGGACGCGCATCGACAAGGCGCGAGGAGGACCGAGAA 5360

RESULT 10
AX427938
LOCUS AX427938 9785 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 176 from Patent WO0232943.
ACCESSION AX427938
VERSION AX427938.1 GI:21538025
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for genetic immunization
JOURNAL Patent: WO 0232943-A 176 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
Location/Qualifiers
1..9785 /organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 2360 a 2843 c 2731 g 1851 t
ORIGIN
Query Match 80.2%; Score 1970; DB 6; Length 9785;
Best Local Similarity 90.9%; Pred. No. 1.7e-197;
Matches 2174; Conservative 0; Mismatches 185; Indels 33; Gaps 6;

QY 14 TGGCGGAGCGCATGAGCCAGGCGCACCCAGCGCAACATCTGTATGCGAGCGCAGCAACTTCA 73
Db 2967 TGGCGGAGCGCATGAGCCAGGCTGCGAGCAGCACACCATCATGATGCGCGGCGCACTTCA 3026
QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTTCACTGCGGCAAGAGGCGGCACATCGCCCGCA 133
Db 3027 AGGCGCAAGCG---CATCAAGTGTCTCAACTGCGGCAAGAGGCGGCACCTTGGCCCGCA 3083
QY 134 ACTGCGCGCGCCCGCGCAAGAGGCGCTGCTGGAAGTGGGCAAGAGGCGGCACACAGATGA 193
Db 3084 ACTGCGCGCGCTTGGCAAGAGGCGCTGCTGGAAGTGGGCAAGAGGCGGCACACAGATGA 3143
QY 194 AGGACTGACCGAGCGCCAGGCGCAACTTCTTCCGCGAGGAGCTGGCTTCCCGCAGGCA 253
Db 3144 AGGACTGACCGAGCGGAGCGGCTTA-----TAGGGAAGATCTGGCTTCCACAAAGGA 3197
QY 254 AGGCGCGGAGTTCCTCCAGCGAGCAGAAACCGCGCAACAGCGCCACCGCGCGAGCTGC 313

Db 3198 AGCCAGGGAAATTTCTTCAGACGACAGACAGCCAAACAGCCCAACAGAGAGAGGCTTC 3257
QY 314 AGGTGGCGG- ----CGACAAACCCCGCAGCGAGCGCGCGCGAGCGCCAGGGCA- -- 364
Db 3258 AGTTTGGGAAGACAGACAACTCTCCTCTCAGAGAGAGAGCGGATAGACAAGGAAGCTG 3317
QY 365 ---CCCTGAATCTCCCGCAGATCACCTGTGGAGAGCGCCCTGGTGGAGCATCAAGGTGG 421
Db 3318 TATCTTTCCTTCAGATCACTCTTTGGCAGCAGCCCTCGTCACATAAGATAG 3377
QY 422 GCGCCAGATCAAGAGAGCGCTGCTGGACACCGCGCGGCGAGACACCGTGTCTGGAGAGA 481
Db 3378 GGGCCAGCTGAAGAGGGCCCTTCTAGACACCGCGCGCGAGCAGACCGCTGCTGGAGAGA 3437
QY 482 TGAGCTGCCCGCAAGTGAAGCCCAAGATGATCGCGGCATCGGGGCTTCATCAAG 541
Db 3438 TGAACCTTGCCCGCGCTGGAAGCCCAAGATGATCGCGGCATCGGGGCTTCATCAAG 3497
QY 542 TGGCCAGTACGACCATCTCTGATCGAGATCTGCGGCAAGAGGCGCATCGGCACCGTGC 601
Db 3498 TGGCCAGTACGACCATCTCTGATCGAGATCTGCGGCCACAGGCCATCGGCACCGTGC 3557
QY 602 TGATGGCCCCACCCCGTGAACATCATCGGCGCGAATGCTGACCCAGCTGGGTGCA 661
Db 3558 TGGTGGCCCCACCCCGTGAACATCATCGGCGCGAATGCTGACCCAGATCGGTGCA 3617
QY 662 CCCTGAATTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATGG 721
Db 3618 CCCTGAATTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGGCATGG 3677
QY 722 ACGCCCCAAGGTGAAGCATGTCCTGATCGAGATCTGCGGCAAGAGATCAAGGCCCTGACGCCA 781
Db 3678 ACGCCCCAAGGTGAAGCATGTCCTGATCGAGATCTGCGGCCACAGGCCATCGGCACCGTGC 3737
QY 782 TCTGCGAGGAGATGGAAGAGGGGCAAGATCAACAGATCGGCCCGCGAGAACCCCTACA 841
Db 3738 TCTGCAACGAGATGGAAGAGGGGCAAGATCAACAGATCGGCCCGCGAGAACCCCTACA 3797
QY 842 ACACCCCGCTTTCGCATCAAGAAAGGACAGACACCAAGTGGCGCAAGCTGTGGACT 901
Db 3798 ACACCCCGCTTTCGCATCAAGAAAGGACAGACACCAAGTGGCGCAAGCTGTGGACT 3857
QY 902 TCGCGAGCTGAACACGCGCACCCAGGACTTCTGGAGGTGCAAGCTGGGCAATCCCCACC 961
Db 3858 TCGCGAGCTGAACACGCGCACCCAGGACTTCTGGAGGTGCAAGCTGGGCAATCCCCACC 3917
QY 962 CCGCCGCTGAAGAGAGAGAGAGCTGACCTGCTGGAGCTGGGGAGCGCTACTTCA 1021
Db 3918 CCGCCGCTGAAGAGAGAGAGAGCTGACCTGCTGGAGCTGGGGAGCGCTACTTCA 3977
QY 1022 GCGTGCCCTGGACGAGGACTTCGCAAGTACACCGCTTTCACCATCCCGAGCATCAACA 1081
Db 3978 GCGTGCCCTGGACAGGACTTCGCAAGTACACCGCTTTCACCATCCCGAGCATCAACA 4037
QY 1082 ACAGACCCCGCGCATCCGCTACCACTACCAAGCTGTGCCCGCCAGGCTGGAAGGGCAGCC 1141
Db 4038 ACAGACCCCGCGCATCCGCTACCACTACCAAGCTGTGCCCGCCAGGCTGGAAGGGCAGCC 4097
QY 1142 CCAGCATCTCCAGAGCAGCATGACCAAGATCTTGAGCCCTTCGGGCGCCGCAACCCCG 1201
Db 4098 CCGCCATCTTCCAGTGCAGCATGACCAAGATCTTGAGCCCTTCGGCAGCAAGCAACCCCG 4157
QY 1202 AGATCGTGTACTACCA- ----GGCCCCCTGTACGTGGGCGCAGCTGGAGATCGGCC 1255
Db 4158 ACATCGTGTACTACAGTACATGACCACTGTACGTGGGCGCAGCTGGAGATCGGCC 4217
QY 1256 AGCACCGCGCAAGATCGAGGAGCTGCGCAAGCACTGTGCGCTGGGGTTCACCAACC 1315
Db 4218 AGCACCGCAAGATCGAGGAGCTGCGCAAGCACTGTGCGCTGGGGTTCACCAACC 4277
QY 1316 CCGCAAGAGACACAGAGAGGAGCCCTCTCTGCCCAT- ----CGAGCTGCACCCCG 1369
Db 4278 CCGCAAGAGACACAGAGAGGAGCCCTCTCTGCCCAT- ----CGAGCTGCACCCCG 4337

QY 1370 ACAAGTGGACCTGACGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCGTGAACGACA 1429
Db 4338 ACAAGTGGACCTGACGCCCATCGTGTGCCGAGAGAGAGCTGGACCGTGAACGACA 4397
QY 1430 TCAGAGCTGGTGGGCAAGCTGAACCTGGCCGACGACAGATCTACCCCGGATCAAGGTGC 1489
Db 4398 TCAGAGCTGGTGGGCAAGCTGAACCTGGCCGACGACAGATCTACCCCGGATCAAGGTGC 4457
QY 1490 GCAGCTGTCAAGCTGTCTCGGGGCGCCAAAGCCCTGACCGACATCGTCCCTGACCG 1549
Db 4458 GCAGCTGTCAAGCTGTCTCGGGGCGCCAAAGCCCTGACCGAGGTGTGCCCTGACCG 4517
QY 1550 AGGAGCGCGAGCTGGAGCTGGCCGAGAAACCGAGAGATCTCTGCGGAGCCCGTGCACCGCG 1609
Db 4518 AGGAGCGCGAGCTGGAGCTGGCCGAGAAACCGAGAGATCTCTGAGGAGCCCGTGCACCGCG 4577
QY 1610 TGTACTACGACCCAGCAAGGAGCTGGTGGCCGAGATCTCCAGAGCAGGCGCAGACCACT 1669
Db 4578 TGTACTACGACCCAGCAAGGAGCTGTATGCGCCGAGATCTCCAGAGCAGGCGCAGGCGCAGT 4637
QY 1670 GGACCTACAGATCTTACCAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCAAGA 1729
Db 4638 GGACCTACAGATCTTACCAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 4697
QY 1730 TGCACACCCGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAAGATCGCCA 1789
Db 4698 TGAAGGCGGCCACCAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAAAGATCGCCA 4757
QY 1790 TGGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCGCCCTGCCCATCCAGAAAGAGA 1849
Db 4758 CCGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCAGCTGCCATCCAGAAAGAGA 4817
QY 1850 CTTGGGAGACCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTGC 1909
Db 4818 CTTGGGAGGCTGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTGC 4877
QY 1910 TGAACACCCCGCTTGTGTAAGCTGTGTACAGCTTGGAGAGGAGGCCATCATCGGCG 1969
Db 4878 TGAACACCCCGCTTGTGTAAGCTGTGTACAGCTTGGAGAGGAGGCCATCATCGGCG 4937
QY 1970 CCAGACCTTCTACGTGGACGCGCGCCCAACCGCAGACCAAGATCGGAAAGGCCCGCT 2039
Db 4938 CCGAGACCTTCTACGTGGACGCGCGCCCAACCGCAGACCAAGCTGGGCAAGGCCCGCT 4997
QY 2030 ACSTGACCGACCGGCGCGCAGAAAGTCTGTAGCCTGACCGAGACCAACCAAGAGAGA 2089
Db 4998 ACSTGACCGACCGGCGCGCAGAAAGTGTGTGCCCTTGACCCGACACCAACCAAGAGAGA 5057
QY 2090 CCGAGCTGACGCGCATCTCAGCTGGCCCTGCAGGACAGCGCAGGAGTGAACATCTGA 2149
Db 5058 CCGAGCTGACGCGCATCTCAGCTGGCCCTGCAGGACAGCGGCTGGAGGTGAACATCTGA 5117
QY 2150 CCACAGCCAGTACGCGCTTGGGCTATCTCCAGGCCCAAGCCCGCAGAGAGGAGCGAGC 2209
Db 5118 CCGACAGCCAGTACGCGCTTGGGCTATCTCCAGGCCCAAGCCCGCAGAGAGGAGCGAGC 5177
QY 2210 TGGTGAACAGATCATCGACAGCTGATCAAGAGAGAGAGTGTACTGAGCTGGGTGC 2269
Db 5178 TGGTGAACAGATCATCGACAGCTGATCAAGAGAGAGAGTGTACTGAGCTGGGTGC 5237
QY 2270 CCGCCACAGGCGCATCGGCGCAACGAGCAGATCGACAGCTGGTGAAGAGGCGCATCC 2329
Db 5238 CCGCCACAGGCGCATCGGCGCAACGAGCAGGAGGAGGCGGCTGGAGGCGCGCATCC 5297
QY 2330 GCAAGTGTCTTCTTCCCTGGACGCGATCGATGGCGGCGATCGTGTACTACAGTA 2381
Db 5298 GCAAGTGTCTTCTTCCCTGGACGCGATCGACAAAGGCCCGAGGAGGAGCAGAGAA 5349
RESULT 11
AX427933
LOCUS AX427933 9167 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 171 from Patent WO0232943.
ACCESSION AX427933
VERSION AX427933.1 GI:21538020
KEYWORDS synthetic construct.
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Huang,Y. and Nabel,G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 171 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source location/Qualifiers
1. .9167
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2225 a 2668 c 2517 g 1757 t
ORIGIN

Query Match 80.0%; Score 1966.6; DB 6; Length 9167;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

Qy 14 TGGCGGAGCCATGAGCCAGG---CCACCAGCGCCACACCTCTGATGAGCGCGCAACT 70
Dy 2964 TGGCGGAGCCATGAGCCAGGTTGACACAGCGCCACCATCATGATGAGCGCGCAACT 3023
Qy 71 TCAAGGGCCCCAAGGGCATCATCAAGTGTCTCAACTCGGGCAAGGAGGGCCACATCGCCC 130
Dy 3024 TCCGCAACCGCGCAAGTGTGAAGTGTCTCAACTCGGGCAAGGAGGGCCACATCGCCC 3083
Qy 131 GCAACTGCGCGCGCCCGCGCAAGAGGGTGTCTGGAAGTGGCGCAAGGAGGGCCACACAGA 190
Dy 3084 GCAACTGCGCGCGCCCGCGCAAGAGGGTGTCTGGAAGTGGCGCAAGGAGGGCCACACAGA 3143
Qy 191 TGAAGGACTGACCGAGCGCGCCAGCGCAACTTCTTCCGGGAGGAGCTTGGCTTCCCCCAGG 250
Dy 3144 TGAAGGACTGACCGAGCGCGCGAGCGCTAA-----TAGGGAAGATCTGGCCCTCCACAA 3197
Qy 251 GCAAGGGCCGGAGTTCGCCAGCGAGGACGAGCGCCGCAAGCGCCACACGAGCGCGAGC 310
Dy 3198 GGAAGGCGAGGAATTTCTTCAGAGGAGACGAGCGCCGCAAGCGCCACACGAGAGAGC 3257
Qy 311 TGCAGGTGCGCGG-----CGACAAACCGCGCGAGCGCGCGCGCGCGCGCGCGCGCA 364
Dy 3258 TTCAGGTTTGGGGAAGAGACAACTCCCTCTCAGAGAGCAGAGCGCGATAGACAAAGAA 3317
Qy 365 -----CCCTGAATTCGCCAGATCACCTGTGGAGCGCGCGCGCGCGCGCGCGCGCG 418
Dy 3318 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCGCTCGTCAACAATAAGA 3377
Qy 419 TGGCGGCGCAGATCAAGGAGCGCTGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 478
Dy 3378 TAGGGGCGCAGCTGAGGAGCGCCCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCG 3437
Qy 479 AGATGAGCTTCCCGGCAAGTGGAGCGCCCAAGATGATGCGCGCGATCGCGCGCTTCATCA 538
Dy 3438 AGATGAACCTGCCCGCGCGTGGAGCCCAAGATGATGCGCGCGATGATGCGCGCGCTTCATCA 3497
Qy 539 AGGTGCGCAGTACGACAGATCTGTATGATGATGATGATGATGATGATGATGATGATGAT 598
Dy 3498 AGGTGCGCAGTACGACAGATCTGTATGATGATGATGATGATGATGATGATGATGATGAT 598
Qy 599 TGCTGATCGGCGCCCGCGCGTGAACATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 658
Dy 3558 TGCTGTTGGGCGCCCGCGCGTGAACATCATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3617
Qy 659 GCACCTGAACTTCCCGATCAGCGCCCGCGCGTGAACATCATCGCGCGCGCGCGCGCGCG 718
Dy 3618 GCACCTGAACTTCCCGATCAGCGCCCGCGCGTGAACATCATCGCGCGCGCGCGCGCGCG 3677

Qy 719 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGAGAGATCAAGGCCCTTGACCG 778
Dy 3678 TGGAGCGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGAGAGATCAAGGCCCTTGTTGG 3737
Qy 779 CCATCTGCGAGGAGATGGAGAGGAGGCGACATCACCAGATCGGCCCGCGAGACCCCT 838
Dy 3738 AGATCTGCGAGGATGGAGAGGAGGCGACATCAGCAAGATCGGCCCGCGAGACCCCT 3797
Qy 839 ACAACACCCCGTGTTCGCCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
Dy 3798 ACAACACCCCGTGTTCGCCCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 3857
Qy 899 ACTTCCGCGAGTGAACAAGCGCACCCAGGACTTCTGGAGAGTGCAGCTGGCGATCCCC 958
Dy 3858 ACTTCCGCGAGTGAACAAGCGCACCCAGGACTTCTGGAGAGTGCAGCTGGCGATCCCC 3917
Qy 959 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGTCTGGACGTGGCGACGCGCTACT 1018
Dy 3918 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGTCTGGACGTGGCGACGCGCTACT 3977
Qy 1019 TCAGCTGCCCTTGACGAGGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA 1078
Dy 3978 TCAGCTGCCCTTGACGAGGAGACTTCCGCAAGTACACCGCCTTCACCATCCCCAGCATCA 4037
Qy 1079 ACAAGGACCCCGCGCATCCGCTACCAAGTACACAGTGTGCCCCAGGGCTGGAAGGCA 1138
Dy 4038 ACAAGGACCCCGCGCATCCGCTACCAAGTACACAGTGTGCCCCAGGGCTGGAAGGCA 4097
Qy 1139 GCGCCAGATCTCCAGAGCAGTACCAAGATCCTGGAGCCCTTCGCGCGCCGCAACC 1198
Dy 4098 GCGCCAGATCTCCAGTGCAGATGACCAAGATCCTGGAGCCCTTCGCGCAAGCAAGACC 4157
Qy 1199 CCGAGATCGTATCTACCA-----GGCCCGCCCTCTAGCTGGCGAGCGACCTTGGAGATCG 1252
Dy 4158 CCGAGATCGTATCTACCAAGTACATGAGCAGCTCTAGCTGGCGAGCGACCTTGGAGATCG 4217
Qy 1253 GCGAGACCGCGCGCAAGTTCGAGGAGTGCAGACACTGCTGCGCTGGGCTTTCACCA 1312
Dy 4218 GCGAGACCGCGCGCAAGTTCGAGGAGTGCAGACACTGCTGCGCTGGGCTTTCACCA 4277
Qy 1313 CCGCGCAAGAGACCGAGAGGAGCGCCCTTCTTCCGCGCAT-----CGAGCTGCGACC 1366
Dy 4278 CCGCGCAAGAGAGACCGAGAGGAGCGCCCTTCTTCTGGATGGCTACGAGCTGCGACC 4337
Qy 1367 CCGACAAGTGGACCGTGCAGCGCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACG 1426
Dy 4338 CCGACAAGTGGACCGTGCAGCGCATCGTGCAGCGCATCGTGCAGAGAGAGCTGGACCGTGAACG 4397
Qy 1427 ACATCCAGAGTGTGGGCAAGCTGAAGTGGGCGAGCGAGATCTACCGCGCATCAAGG 1486
Dy 4398 ACATCCAGAGTGTGGGCAAGCTGAAGTGGGCGAGCGAGATCTACCGCGCATCAAGG 4457
Qy 1487 TCGCGCAGCTGTGAAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1546
Dy 4458 TCGCGCAGCTGTGAAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4517
Qy 1547 CCGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGCGCGCGCGCGCG 1606
Dy 4518 CCGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTGCGCGAGCGCGCGCGCGCG 4577
Qy 1607 GCGTGTACTACGACCGCGAGAGGAGCTGGTGGCGCGAGATCCAGAGAGGCGCGACGAC 1666
Dy 4578 GCGTGTACTACGACCGCGAGAGGAGCTGGTGGCGCGAGATCCAGAGAGGCGCGACGAC 4637
Qy 1667 AGTGACCTTACAGATCTACCGAGGAGCGCTTCAAGAACTTGAAGACCGCGCGAGTACGCCA 1726
Dy 4638 AGTGACCTTACAGATCTACCGAGGAGCGCTTCAAGAACTTGAAGACCGCGCGAGTACGCCA 4697
Qy 1727 AGATGCGCACCGCGCGCACACCAAGAGCTGAAGAGCTGACGAGGCGCGTGCAGAGATCG 1786
Dy 4698 GCATGAAGGCGCGCGCACACCAAGAGCTGAAGAGCTGACGAGGCGCGTGCAGAGATCG 4757
Qy 1787 CCATGGAGAGCATCGTGTATCTGGGCAAGACCGCGCGCGCGCGCGCGCGCGCGCGCG 1846


```
Db 4758 CCACCGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCATCCAGAAGG 4817
QY 1847 AGACCTGGGAGACCTGGTGACCGACTACTGGCAGGCCACTTGATCCCGAGTGGGAGT 1906
Db 4818 AGACCTGGGAGGCTGGTGACCGAGTACTGGCAGGCCACTTGATCCCGAGTGGGAGT 4877
QY 1907 TCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGCCCATCATCG 1966
Db 4878 TCGTGAACACCCCCCTGGTGAAGCTGTGGTACCAGCTGGGAGAGGCCCATCATCG 4937
QY 1967 GCGCCGAGACCTTCTAGCTGGAGCGGCGCCAGAGATCGTGACCTTGACCGAGACCACCAACAGA 2086
Db 4938 GCGCCGAGACCTTCTAGCTGGAGCGGCGCCAGAGTGTGGCCCTGACCGACACCAACCAAGC 5057
QY 2087 GCTACGCTGACCGAGCCGCGGCGCCAGAGATCGTGACCTTGACCGAGACCACCAACAGA 2086
Db 4998 GCTACGCTGACCGAGCCGCGGCGCCAGAGTGTGGCCCTGACCGACACCAACCAAGC 5057
QY 2087 AGACCCAGCTGACGAGCCATCCAGCTGGCCCTGACGAGACGCGGAGGTGAACATCG 2146
Db 5058 AGACCCAGCTGACGAGCCATCCAGCTGGCCCTGACGAGACGCGGAGGTGAACATCG 5117
QY 2147 TGACCCGACAGCAGTAGTCCGCTGGGCATCATCCAGGCCCGCCAGCAGGAGAGCG 2206
Db 5118 TGACCCGACAGCAGTAGTCCGCTGGGCATCATCCAGGCCCGCCAGCAGGAGAGCG 5177
QY 2207 AGCTGGTGAACAGATCATCGACGAGCTGATCAAGAGGAGAGGTTACTGAGCTGGG 2266
Db 5178 AGCTGGTGAACAGATCATCGACGAGCTGATCAAGAGGAGAGGTTACTGAGCTGGG 5237
QY 2267 TGCCCGCCCAACAGGCGATCGGGGCAACGAGCAGATCGACAAGCTGGTGAAGAGGCA 2326
Db 5238 TGCCCGCCCAACAGGCGATCGGGGCAACGAGCAGTGGAGCGCTGGTGAAGAGGCA 5297
QY 2327 TCAGCAAGTGTCTTCTCGAGCGGATCGATGGCGGCATCGTGATCTACCAAGTA 2381
Db 5298 TCAGCAAGTGTCTTCTCGAGCGGATCGACAAGGCCCGGAGGAGGAGGCAAGAA 5352

RESULT 12
AX427928
LOCUS AX427928 9170 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 166 from Patent WO0232943.
ACCESSION AX427928
VERSION AX427928.1 GI:21538015
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Huang, Y. and Nabel, G. J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
JOURNAL genetic immunization
PATENT Patent: WO 0232943-A 166 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
1. 9170
/organism="synthetic construct"
/db_xref="taxon:32830"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 225 a 2669 c 2519 g 1757 t
ORIGIN
```

```
Query Match 80.0%; Score 1966.6; DB 6; Length 9170;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCCGAGGCGCATGAGCAGG---CCACCAGCGCCCAACATCTCTGATCGACGAGCAACT 70
Db 2967 TGGCCGAGGCGCATGAGCAGGCTGACCCAGAGCGCCACCATCATGATGACGCGGGCACT 3026
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTTCAACTGCGGCAAGGAGGCGGCACATCGCCC 130
```

```
Db 3027 TCAGCAACACCGAGCGCAAGATCGTGAAGTGTCTTCAACTGCGCAAGGAGGCGGCACACCCCC 3086
QY 131 GCAACTCCCGCGCCCGCCCAAGAGAGGCTGTGGAAGTGGCGCAAGGAGGCGGCACACAGA 190
Db 3087 GCAACTCCCGCGCCCGCCCAAGAGAGGCTGTGGAAGTGGCGCAAGGAGGCGGCACACAGA 3146
QY 191 TGAAGGAGCTCACCGAGCGCGCAGCCCAACTTCTTCCGCGAGGAGCTGGCTTCCCTCCCGCAGG 250
Db 3147 TGAAGGAGCTCACCGAGCGCGCAGCCCAACTTCTTCCGCGAGGAGCTGGCTTCCCTCCCGCAGG 3200
QY 251 GCAAGGCGCCCGAGTTCGCCAGCGAGAGCAACCGCCCAACAGCCCGCCAGCGCGCGCAGGCA 310
Db 3201 GGAAGGCGCAGGAGTTCCTTTCAGAGCAGACAGAGCAACAGCGCCCGCCAGAGAGAGC 3260
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGCGCGCAGGCGCA 364
Db 3261 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGCGCGCGCGCGCGCGCAGGCGCA 3320
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGTGAGCATCAAGG 418
Db 3321 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCGACCCCTGTGACATTAAGA 3380
QY 419 TGGCGGCCAGATCAAGGAGGCGCCCTGTGTGACACCGCGCGCGCGCAGCACCCGTGCTGAGG 478
Db 3381 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGCGCAGCACCCGTGCTGAGG 3440
QY 479 AGATGAGCTTCCCGCGCAAGTGGAGCCCAAGATGATGCGGCGCATCGCGGCTTCATCA 538
Db 3441 AGATGAACTTCCCGCGCGCGCTGGAAAGCCCAAGATGATGCGGCGCATCGCGGCTTCATCA 3500
QY 539 AGGTGCGCCAGTAGTACGACAGATCCTGTGATCGAGATCTCGGCAAGAGGCCATCGGCACCG 598
Db 3501 AGGTGCGCCAGTAGTACGACAGATCCTGTGATCGAGATCTCGGCGCCACAAGGCCATCGGCACCG 3560
QY 599 TGTGATGCGCGCCACCGCCGTGAACATCATTCGCGCGCAACATGCTGACCGAGCTGGGCT 658
Db 3561 TGTGTTGGCGCCCGCCCGCTGAAACATCATTCGCGCGCAACCTGCTGACCGCAGATCGGCT 3620
QY 659 GCACCTGAACTTCCCGCATCAGCCCGCATCGAGACCGTGGCGCTGAAGCTGAAGCGCGCA 718
Db 3621 GCACCTGAACTTCCCGCATCAGCCCGCATCGAGACCGTGGCGCTGAAGCTGAAGCGCGCA 3680
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGCTGG 3740
QY 779 CCATCTGCGAGGAGATGGAGAGGCGCAAGATCAACAGATCGGCCCGCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAGAGGCGCAAGATCAGCAAGATCGGCCCGCGAGAACCCCT 3800
QY 839 ACACACCGCCCGTTCGCCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTGG 898
Db 3801 ACAACACCGCCCGTTCGCCCATCAAGAGAGGAGACACCAAGTGGCGCAAGCTGGTGG 3860
QY 899 ACTTCCCGGAGCTGAACAAGCGACCCAGGACTTCTTGGAGTGCAGTGGGCGATCCCGC 958
Db 3861 ACTTCCCGGAGCTGAACAAGCGACCCAGGACTTCTTGGAGTGCAGTGGGCGATCCCGC 3920
QY 959 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGCTGTGGAGTGGGCGACGCGCTACT 1018
Db 3921 ACCCGCGCGCTGAAGAAGAGAGCGGTGACCGCTGTGGAGTGGGCGACGCGCTACT 3980
QY 1019 TCAGCTGCGCCCTGGAGAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 1078
Db 3981 TCAGCTGCGCCCTGGAGAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCCAGCATCA 4040
QY 1079 ACAACAGAGACCCCGGATCCGCTACCAAGTACACAGTGTGCGCCCGAGGCTGGAGGCA 1138
Db 4041 ACAACAGAGACCCCGGATCCGCTACCAAGTACACAGTGTGCGCCCGAGGCTGGAGGCA 4100
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCGCTTCCCGCGCGCGCAACC 1198
```


Db 4101 GCGCGCCATCTTCCAGTGCAGCATGACCAAGATCCTTGAGCCCTTCGCGAAGCAGAAC 4160
Qy 1199 CCGAGATGCTGATCTACCA-----GGCCGCCCTGTAGTGGGACGACCTGGAGATCG 1252
Db 4161 CCGACATGCTGATCTACCAAGTACATGGACACCTGTAGTGGGACGACCTGGAGATCG 4220
Qy 1253 GCGACACCGCGGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCGTGGGCTTCAACA 1312
Db 4221 GCGACACCGCGGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGCGTGGGCTTCAACA 4280
Qy 1313 CCGCGCAAGAACACACAGAGAGCGCCCTTCTCTGCCAT-----CGAGCTGCACC 1366
Db 4281 CCGCGCAAGAACACACAGAGAGCGCCCTTCTCTGGATGGGCTACGAGCTGCACC 4340
Qy 1367 CCGACAAGTGGACCGTGCAGCGCCATCGAGCTGCCGCGAAGAGAGCTGGACCGTGAACG 1426
Db 4341 CCGACAAGTGGACCGTGCAGCGCCATCGTGTCTGCCGAGAGAGACGCTGGACCGTGAACG 4400
Qy 1427 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGGCATCAAG 1486
Db 4401 ACATCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGAGCCAGATCTACCGCGCATCAAG 4460
Qy 1487 TGGCCGAGCTGTCAAGCTGTCTGCGCGCGCCAGGCGCTTACCGACATCTGCGCCCTGA 1546
Db 4461 TGGCCGAGCTGTCAAGCTGTCTGCGCGCGCCAGGCGCTTACCGAGGTGTGCGCCCTGA 4520
Qy 1547 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGCGGAGCGCGTGCACG 1606
Db 4521 CCGAGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCCTGAGAGCGCGTGCACG 4580
Qy 1607 GCGTGTACTAGACCCAGCAAGACCTGTGGCGCGAGATCCAGAAAGCAGGCGCCACGACC 1666
Db 4581 GCGTGTACTAGACCCAGCAAGACCTGTATCGCGGAGATCCAGAAAGCAGGCGCGCCACG 4640
Qy 1667 AGTGGACCTACACAGATCTACCAAGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 4641 AGTGGACCTACCAAGATCTACCAAGAGCGCTTCAAGAACCTTGAAGACCGGCAAGTACGCC 4700
Qy 1727 AGATGCGCAACCGCCACACCAACAGCTGTGAAGCAGCTGACCGAGGCGCTGCAGAAAGTCG 1786
Db 4701 GCATGAAGGGCGCCACACCAACAGCTGAAGCAGCTGACCGAGGCGCTGCAGAAAGTCG 4760
Qy 1787 CCATGGAGAGATCGTGTATCTGGGCGCAAGACCCCGAGTTCCGGCTGCCCATCCAGAAGG 1846
Db 4761 CCACGAGAGATCGTGTATCTGGGCGCAAGACCCCGAGTTCAAGCTGCCCATCCAGAAGG 4820
Qy 1847 AGACCTGGGAGACCTGTGTGAGCAGCTTACTGGCAGGCGCACTGTGATCCCGAGTGGGAGT 1906
Db 4821 AGACCTGGGAGGCTGTGTGAGCAGCTTACTGGCAGGCGCACTGTGATCCCGAGTGGGAGT 4880
Qy 1907 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGCTGGAGAAGGAGCCCATCATCG 1966
Db 4881 TCGTGAACACCCCGCCCTGGTGAAGCTGTGTACCAAGCTGGAGAAGGAGCCCATCATCG 4940
Qy 1967 GCGCGGAGACCTTCTAGCTGGAGCGCGCGCCACACCGGAGACCAAGATGGCAAGGCGG 2026
Db 4941 GCGCGGAGACCTTCTAGCTGGAGCGCGCGCCACACCGGAGACCAAGCTGGCAAGGCGG 5000
Qy 2027 GCTACGTGACCGACCGGCGCGGCGAGAGATCGTGAAGCTGTGTGAGCTGGAGAAGGAGCC 2086
Db 5001 GCTACGTGACCGACCGGCGCGGCGAGAGGTGTGTGAGCTGGAGGAGGAGGAGGAGG 5060
Qy 2087 AGACCGAGCTGCAGGCGCATCAGCTGGCCCTGCAGGACAGCGGAGGAGGTGAACATCG 2146
Db 5061 AGACCGAGCTGCAGGCGCATCAGCTGGCCCTGCAGGACAGCGGCGCTGGAGGTGAACATCG 5120
Qy 2147 TGACCGACAGCAGTACGCGCTGGGCGATCATCCAGGCGCCAGCGCGCAAGAGCGAGAGCG 2206
Db 5121 TGACCGACAGCAGTACGCGCTGGGCGATCATCCAGGCGCCAGCGCGCAAGAGCGAGAGCG 5180
Qy 2207 AGCTGTGTAACCCAGATCATCGACAGCTGATCAAGAGGAGAGAGGTGTACCTGAGCTGGG 2266
Db 5181 AGCTGTGAGCCAGATCATCGACAGCTGATCAAGAGGAGAGAGGTGTACCTGAGCTGGG 5240

Qy 2267 TGCCCGCCACAAAGGGCATCGCGGCAAGCAGCAGATCGCAAGCTGGTGACCAAGGGCA 2326
Db 5241 TGCCCGCCACAAAGGGCATCGCGGCAAGCAGCAGATCGCAAGCTGGTGACCGCGGCA 5300
Qy 2327 TCCGCAAGTGTCTTCTCTGGACGCGCATCGATGGCGGATCTGATCTACCAAGTA 2381
Db 5301 TCCGCAAGTGTCTTCTCTGGACGCGCATCGCAAGGCGCCAGGAGGACGACAGAA 5355

RESULT 13
AX427935
LOCUS AX427935 9782 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 173 from Patent WO0232943.
ACCESSION AX427935
VERSION AX427935.1 GI:21538022
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Huang,Y. and Nabel,G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 173 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source Location/Qualifiers
1..9782
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"
BASE COUNT 2339 a 2895 c 2719 g 1829 t
ORIGIN

Query Match 80.0%; Score 1966.6; DB 6; Length 9782;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

Qy 14 TGGCGGAGGCGCATGAGCCAGG---CCACCAGGCGCAACATCTGTATGATGAGCGCAGCACT 70
Db 3579 TGGCGGAGGCGCATGAGCCAGGCGCAACAGCGCCACCATCATGATGATGAGCGGCGCACT 3638
Qy 71 TCAAGGGGCGGCGCATCATCAAGTGTCTTCAAGTGTGGGCAAGGAGGCGGCACATCGGCC 130
Db 3639 TCCGCAACACGCGCAAGATCGTGAAGTCTTCACTGGCGCAAGGAGGCGGCACACCGGCC 3698
Qy 131 GCACTGCGGCGCGCCCGCGCAAGAGGCTGTGTGAAGTGGCGCAAGAGGCGGCACACAGA 190
Db 3699 GCACTGCGGCGCGCCCGCGCAAGAGGCTGTGTGAAGTGGCGCAAGAGGCGGCACACAGA 3758
Qy 191 TGAAGGACTGCACGAGCGCGCCAGCCCACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGG 250
Db 3759 TGAAGGACTGCACGAGCGCGCCAGCCCACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGG 3812
Qy 251 GCAAGGCGCGGAGTTTCCCGAGCGAGCAACCGCGCAACAGCCCGCCACCGCCGCGAGC 310
Db 3813 GGAAGCGCAGGAGTTTCTTCAAGCAGACAGCAGCCAGCCCGCCACCGCCAGAGAGAGC 3872
Qy 311 TGCAGTGTGCGGG-----CGACAACCGCCGAGCGCGCGCGCGCGCGCGCGCGCGCA 364
Db 3873 TTCAGTTTGGGGAGAGACAACTCCCTCTCAGAAAGCAGGAGCGCCATAGACAAGGAA 3932
Qy 365 -----CCCTGAACCTTCCCGAGATCACCCTGTGGCAGCGCCCTTGTGTGAGCATCAAG 418
Db 3933 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTTGGCAGCGACCCCTCGTCAATAAAGA 3992
Qy 419 TGGCGCGCGAGATCAAGAGGAGCCCTGTGTGACACCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 3993 TAGGGCGCGAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCG 4052
Qy 479 AGATGAGCTGCCCGGCAAGTGGAAAGCCCAAGATCATCGCGCGCATCGCGCGCTTCATCA 538
Db 4053 AGATGAACCTGCCCGCGCGCTTGAAGGCCCAAGATCATCGCGCGCATCGCGCGCTTCATCA 4112

BASE COUNT 2339 a 2897 c 2723 g 1824 t
ORIGIN /note="plasmid pVR1012x/s containing HIV genes"

Query Match 80.0%; Score 1966.6; DB 6; Length 9783;
Best Local Similarity 90.7%; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCGAGGCGCATGAGCAGG--CCACCAGCGCCCAACATCTGTGATGAGCGCAGCAACT 70
DB 2964 TGGCGAGGCGCATGAGCAGGTCACCAAGCGCCCAACATGATGAGCGCGCAACT 3023
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGCGCCACATCGGCC 130
DB 3024 TCCGCAACAGCGCAAGATCGTGAAGTCTTCAACTGCGGCAAGAGGCGCCACACCGGCC 3083
QY 131 GCAACTGCGCGCCCCCGCCGCAAGAGGCTGTGGAATGCGGCAAGAGGCGCCACCAAGA 190
DB 3084 GCAACTGCGCGCCCCCGCCGCAAGAGGCTGTGGAATGCGGCAAGAGGCGCCACCAAGA 3143
QY 191 TGAAGGACTGCACCGAGCGCCAGGCAACTTCTTCCGCGAGGACCTGGCCTTCCCGCAGG 250
DB 3144 TGAAGGACTGCACCGAGCGCAGGCTAA-----TAGGAGATCTGGCCTTCCCAAG 3197
QY 251 GCAAGGCCGCGAGTTCCCGCAGCAGAGAACCGCGCCCAAGAGCCCGCCACAGCGCGAGC 310
DB 3198 GGAAGGCCAGGGAATTTTCTTCAAGCAGACAGAGCCCAAGCGCCCGCCACAGAGAGAGC 3257
QY 311 TGCAGTCCGCGG-----CGACAAACCCCGAGCGGCGCGGCGCCGAGCGCCAGGCGCA 364
DB 3258 TTCAGGTTTGGGGAAGAGAACCAACTCTCTCAGAAAGCAGGAGCGCATGAGCAAGGAA 3317
QY 365 -----CCCTGAATCTCCCGCAGATCACCTGTGCGACGCGCCCTTGTGAGCATCAGG 418
DB 3318 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGCAGCCCTCTGTCACAAATAAGA 3377
QY 419 TGGCGGCGCAGATCAAGAGGCGCCTGTGTGACACCGCGCGCAGACACCGTGTGGAGG 478
DB 3378 TAGGGGCGCAGCTGAAGAGGCGCCTTCTAGACACCGGCGCGCAGACACCGTGTGGAGG 3437
QY 479 AGATGAGCTGCCCGGCAAGTTGGAAGCCCAAGATGATGCGGCGCATCGCGCGCTTCATCA 538
DB 3438 AGATGAACTTCCCGGCGCGCTGGAAGCCCAAGATGATGCGGCGCATCGCGCGCTTCATCA 3497
QY 539 AGGTGCGCAGTACGACAGATCTCTGATCGAGATCTGCGGCAAGAGGCCATCGCACCG 598
DB 3498 AGGTGGCCAGTACGACAGATCTCTGATGASATCTGCGGCGCACAGGCCATCGCACCG 3557
QY 599 TGCTGATGCGGCGCCACCGCGGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCT 658
DB 3558 TGCTGGTGGCGCCACCGCGGTGAACATCATCGCGCGCAACATGCTGACCGAGCTGGGCT 3617
QY 659 GCACCTGAACTTCCCATCAGCCCGCATCGAGACCGTGCCTGTAAGTGAAGCGCGGCA 718
DB 3618 GCACCTGAACTTCCCATCAGCCCGCATCGAGACCGTGCCTGTAAGTGAAGCGCGGCA 3677
QY 719 TGAAGCGCCCAAGGTGAAGCAGTGGCGCCTGACCGGAGGAGAAGATCAAGSCCTGACCG 778
DB 3678 TGAAGCGCCCAAGGTGAAGCAGTGGCGCCTGACCGGAGGAGAAGATCAAGSCCTGATGG 3737
QY 779 CCATCTCGGAGGAGATGGAGAGGCGCAGATCATCAAGATCGCGCGCCCGAGAACCCCT 838
DB 3738 AGATCTGCACCGAGATGGAGAGGAGGCAAGATCAGCAAGATCGCGCGCCCGAGAACCCCT 3797
QY 839 ACAACACCCCGTCTCGCCATCAAGAGAGGACACGACCAAGTGGCGCAAGTGGTGG 898
DB 3798 ACAACACCCCGTCTCGCCATCAAGAGAGGAGACGACCAAGTGGCGCAAGTGGTGG 3857
QY 899 ACTTCCGAGAGTGAACAAGCGCACCGAGGACTTCTGGAGGTGACGTGGGCATCCCC 958
DB 3858 ACTTCCGAGAGTGAACAAGCGCACCGAGGACTTCTGGAGGTGACGTGGGCATCCCC 3917
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGCGTGAACCGTGTGAGCGTGGCGGACGCGCTACT 1018

DB 3918 ACCCGCGCGGCTGAAGCAGAAGAGAGCGTACCGTGTGGACGTGGCGGACGCGCTACT 3977
QY 1019 TCAGGCTGCCCTGGACGAGGACTTCCGCAAGTATACACCGCTTACCATCCCGCAGCATCA 1078
DB 3978 TCAGGCTGCCCTGGACAGGACTTCCGCAAGTATACACCGCTTACCATCCCGCAGCATCA 4037
QY 1079 ACAAGCAGACCCCGGCATCCCGTACCACTACAACTGTGTGCCCGCAGGCGTGAAGGGCA 1138
DB 4038 ACAAGCAGACCCCGGCATCCCGTACCACTACAACTGTGTGCCCGCAGGCGTGAAGGGCA 4097
QY 1139 GCCCGAGCATCTTCCAGAGCAGCATGACCAAGATCTGTGAGCCCTTCCCGCGCCGCAACC 1198
DB 4098 GCCCGCGCATCTTCCAGTGCAGCATGACCAAGATCTGTGAGCCCTTCCCGCAGAGCAACC 4157
QY 1199 CCGAGATCGTGTATTAACCA-----GGCCCCCTGTACTGTGGCGAGCGACCTGGAGATCG 1252
DB 4158 CCGACATCGTGTATTAACCACTATGAGCACTGTACTGTGGCGAGCGACCTGGAGATCG 4217
QY 1253 GCCAGCACCGCGCCCAAGATCGAGGAGCTGCGCAAGCACTGTGTGGCTGGGCTTACCA 1312
DB 4218 GCCAGCACCGCGCCCAAGATCGAGGAGCTGCGCAAGCACTGTGTGGCTGGGCTTACCA 4277
QY 1313 CCCCCGACAAGAACCAAGAGGAGCGCCCTTCTGCCCCAT-----CGAGCTGCACC 1366
DB 4278 CCCCCGACAAGAACCAAGAGGAGCGCCCTTCTGTTGGATGGGCTACGAGCTGCACC 4337
QY 1367 CCGCAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAAGAGAGAGCTGGACCGTGAACG 1426
DB 4338 CCGCAAGTGGACCGTGCAGCCCATCGTGTGCCCGAGAGAGCAGCTGGACCGTGAACG 4397
QY 1427 ACATCCAGAGCTGTGGGCAAGCTGAAGTGGGCGCAGCGAGATCTACCCCGGCATCAAG 1486
DB 4398 ACATCCAGAGCTGTGGGCAAGCTGAAGTGGGCGCAGCGAGATCTACCGCGGCATCAAG 4457
QY 1487 TCGCGACGTGTGAAGTGTGTGCGCGCGCCAAAGCGCTGACCCACACATCTGTGCCCTGA 1546
DB 4458 TCGCGACGTGTGAAGTGTGTGCGCGCGCCAAAGCGCTGACCCAGGTGTGTGCCCTGA 4517
QY 1547 CCGAGAGGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGTGCGGAGCGCGTGCAGC 1606
DB 4518 CCGAGAGGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTGTGAGGAGCGCGTGCAGC 4577
QY 1607 CGGTGTACTACGACCCCGCAGGAGCTGTGTGGCGCAGATCCAGAGAGGCGGCGCAGC 1666
DB 4578 CGGTGTACTACGACCCCGCAGGAGCTGTGTGGCGCAGATCCAGAGAGGCGGCGCAGC 4637
QY 1667 AGTGACCTTACAGATCTACGAGGAGCGCTTCAAGAACTGAAGACCGGCAAGTACGCCA 1726
DB 4638 AGTGACCTTACAGATCTACGAGGAGCGCTTCAAGAACTGAAGACCGGCAAGTACGCCA 4697
QY 1727 AGATGCGGACCGCGCCACACCAAGCAGTGAAGCAGCTGACCGAGGCGCGTGCAGAGATCG 1786
DB 4698 GCATGAAGGCGCGCCACACCAAGCAGTGAAGCAGCTGACCGAGGCGCGTGCAGAGATCG 4757
QY 1787 CCATGAGAGAGCATGTGATCTGGGCAAGACCCCGCAAGTTCGCGCTGCCCATCCAGAGG 1846
DB 4758 CCACCGAGAGCATGTGATCTGGGCAAGACCCCGCAAGTTCGCGCTGCCCATCCAGAGG 4817
QY 1847 AGACTGGGAGACCTGTGTGGACCGCACTACTGTGCGAGCGCCACCTGGATCCCGAGTGGAGT 1906
DB 4818 AGACTGGGAGCGCTGTGTGGACCGCACTACTGTGCGAGCGCCACCTGGATCCCGAGTGGAGT 4877
QY 1907 TCGTGAACACCCCGCTTGTGTGAAGCTGTGTACAGTGGAGAGGAGCGCCATCATCG 1966
DB 4878 TCGTGAACACCCCGCTTGTGTGAAGCTGTGTACAGTGGAGAGGAGCGCCATCATCG 4937
QY 1967 GCGCGAGAGCTTACGTGAGCGCGCCCGCAAGCAGCAGACCAAGATCGCGCAGGCGCG 2026
DB 4938 GCGCGAGAGCTTACGTGAGCGCGCGCCCGCAAGCAGCAGACCAAGTGGGCGAAGGCGCG 4997
QY 2027 GCTACGTGACCGACCGGCGCGCAGAGATCTGTGAGCTGACCGAGACCAACCAAGCA 2086

Db 4998 GCTACGTGACCGACCGCGCGCCAGAGGTGGTGCCTCTCAGACACACCAACAGCA 5057
Qy 2087 AGACCGAGCTGACGAGCCATCCAGCTGCGCTGCAGGACAGCGCAGAGTGAACATCG 2146
Db 5058 AGACCGAGCTGACGAGCCATCCAGCTGCGCTGCAGGACAGCGCCTGGAGGTGAACATCG 5117
Qy 2147 TGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCGCCAGCCGACAAAGAGGAGAGCG 2206
Db 5118 TGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCGCCAGCCGACAAAGAGGAGAGCG 5177
Qy 2207 AGCTGGTGAACAGATCATCGAGAGCTGATCAAGAGGAGAGAGTGTACTGAGCTGGG 2266
Db 5178 AGCTGGTGAACAGATCATCGAGAGCTGATCAAGAGGAGAGAGTGTACTGAGCTGGG 5237
Qy 2267 TGCCCGCCCAAGAGGATCGGGGCAACGACGATCGCAAGCTGCTGAGCAAGGCA 2326
Db 5238 TGCCCGCCCAAGAGGATCGGGGCAACGACGATCGCAAGCTGCTGAGCGCCGCA 5297
Qy 2327 TCCGCAAGGTGCTTCTTCCAGCGGCATCGATGCGCGCATCGTGTATCAACAGTA 2381
Db 5298 TCCGCAAGGTGCTTCTTCCAGCGGCATCGCAAGGCGCCAGGAGGACAGCAAGAA 5352

RESULT 15
AX427932 AX427932 9792 bp DNA linear PAT 20-JUN-2002
LOCUS Sequence 170 from Patent WO0232943.
DEFINITION AX427932
ACCESSION AX427932
VERSION AX427932.1 GI:21538019
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.

REFERENCE
1
AUTHORS Huang, Y. and Nabel, G.J.
TITLE Modifications of hiv env, gag, and pol enhance immunogenicity for
genetic immunization
JOURNAL Patent: WO 0232943-A 170 25-APR-2002;
GOVERNMENT OF THE UNITED STATES (US)
FEATURES
source 1..9792
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="plasmid pVR1012x/s containing HIV genes"

BASE COUNT 2339 a 2898 c 2725 g 1830 t
ORIGIN

Query Match 80.08; Score 1966.6; DB 6; Length 9792;
Best Local Similarity 90.74; Pred. No. 3.9e-197;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

Qy 14 TGGCCGAGGCCATGAGCCAGG---CCACGAGCGCAACATCTGATGAGCGCAGCAACT 70
Db 2967 TGGCCGAGGCCATGAGCCAGGTGACCAACAGCGCCACCATCATGATGAGCGCGGCAACT 3026
Qy 71 TCAGAGGCCCCAAGCGCATCATCAAGTCTTCAACTGCGGAAGGAGGGCCACATCGGCC 130
Db 3027 TCCGCAACGACGCGAAGATCGTGAAGTCTTCAACTGCGGCAAGGAGGGCCACACCGGCC 3086
Qy 131 GCAACTGCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGGAAGGAGGGCCACACAGA 190
Db 3087 GCAACTGCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGGAAGGAGGGCCACACAGA 3146
Qy 191 TGAAGGACTGACCGAGGCGCCAGCCAACTTCTTCCGCGAGGACTGCGCTTCCCGCAGG 250
Db 3147 TGAAGGACTGACCGAGGCGCACAGGCTAA-----TAGGGAAGATCTGGCTTCCACAG 3200
Qy 251 GCAAGGCGCGAGTTCCTCCAGCGAGAGAACCGCGCAACAGCCCCACAGCGCGCGAGC 310
Db 3201 GGAAGGCGCGAGTATTTCTTCAGAGCAGACACAGAGCCCAACAGCCCCACAGAGAGAGC 3260
Qy 311 TGCAGTGGCGGG-----CGACAAACCCCGCAGGAGGCGCGCGCGCGCGCGCGCGCA 364

Db 3261 TTCAAGTTTGGGAAAGAGACAACAACATCCCTCTCAGAAAGCAGGAGCGCATAGACAAGGAA 3320
Qy 365 -----CCCTGAACCTTCCCCAGATCACCTGTGGCAGCGCCCTGGTGGAGCATCAAGG 418
Db 3321 CTGTATCTTTAGCTTCCCTCAGATCACTTTTGGCAGGAGCCCTCTGTACATAAAGA 3380
Qy 419 TGGCGGCCAGATCAAGAGAGCCCTGTGTGACACCGCGCGCGGAGCAGCACCTGTGTGAGG 478
Db 3381 TAGGGGCCAGCTGAAGAGGAGCCCTTCTAGACACCGCGCGCGAGCAGACCTGTGTGAGG 3440
Qy 479 AGATGACCTTCCCGGCAAGTGGAAAGCCCAAGATGATCGGGGCGATCGGGGCTTCATCA 538
Db 3441 AGATGAACTTCCCGGCGCGCTTGGAAAGCCCAAGATGATCGGGGCGATCGGGGCTTCATCA 3500
Qy 539 AGGTGCGCCAGTACGACACAGATCTGTATCGAGATCTGCGGCAAGAAGCCATCGGCACCG 598
Db 3501 AGGTGCGCCAGTACGACACAGATCTGTATCGAGATCTGCGGCAAGAAGCCATCGGCACCG 3560
Qy 599 TGTGTATCGGCGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 3561 TGTGTGTGGGCGCCACCCCGGTGAACATCATCGGCGCGCAACATGCTGACCCAGATCGGCT 3620
Qy 659 GCACCTTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGTGAAGCCCGGCA 718
Db 3621 GCACCTTGAACCTTCCCATCAGCCCATCGAGACCGTGCCTGGAAGTGAAGCCCGGCA 3680
Qy 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAAGATCAAGCCCTGACCG 778
Db 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAAGATCAAGCCCTGTTGG 3740
Qy 779 CCATCTCGAGGAGATGGAGAAAGGAGGCAAGATCAACCAAGATCGGCGCCCGAGAACCCCT 838
Db 3741 AGATCTGCACCGAGATGGAGAAAGGAGGCAAGATCAGCAAGATCGGCGCCCGAGAACCCCT 3800
Qy 839 ACACACCCCGCTTCCCATCAAGAAAGAGCAGACCAAGTGGCGCAAGCTGGTGG 898
Db 3801 ACACACCCCGCTTCCCATCAAGAAAGAGCAGACCAAGTGGCGCAAGCTGGTGG 3860
Qy 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGCAGTGGGCGATCCCCC 958
Db 3861 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGCAGTGGGCGATCCCCC 3920
Qy 959 ACCCGCGCGCTGAAGAAGAAAGAGCGCTGACCGTGTGGAGCTGGCGGCGACGCTACT 1018
Db 3921 ACCCGCGCGCTGAAGAAGAAAGAGCGCTGACCGTGTGGAGCTGGCGGCGACGCTACT 3980
Qy 1019 TCAGCTGCCCTTGACGAGGACTTCCCAAGTACACCGCTTACCATCCCCAGCATCA 1078
Db 3981 TCAGCTGCCCTTGACGAGGACTTCCCAAGTACACCGCTTACCATCCCCAGCATCA 4040
Qy 1079 ACAACGAGACCCCGGCATCCGCTACCAAGTACAAGTGTCTGCCCGAGGGCTGGAGAGGCA 1138
Db 4041 ACAACGAGACCCCGGCATCCGCTACCAAGTACAAGTGTCTGCCCGAGGGCTGGAGAGGCA 4100
Qy 1139 GCGCCAGCATCTTCAGAGAGCATGACCAAGATCTTGAGGCCCTTCCGCGCCCGCAACC 1198
Db 4101 GCGCCAGCATCTTCAGAGTGCAGCATGACCAAGATCTTGAGGCCCTTCCGCGAGCAAGAAC 4160
Qy 1199 CCGAGATCGTGTATCTTACCA-----GGCGCCCTGTAGTGGGCGAGCACCTTGGAGATCG 1252
Db 4161 CCGACATCGTGTATCTTACCAAGTACATGAGCACCTGTAGTGGGCGAGCACCTTGGAGATCG 4220
Qy 1253 GCCAGCAGCGGCCAAGATCGAGGAGCTGCGCAAGCAGCTGTGGCTGGGGCTTTCACCA 1312
Db 4221 GCCAGCAGCGCACCAAGATCGAGGAGCTGCGCAGCAGCTGTGTGGCTGGGGCTTTCACCA 4280
Qy 1313 CCGCGAGCAAGAACACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1366
Db 4281 CCGCGAGCAAGAACACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4340
Qy 1367 CCGCAAGTGGAGCGTGCAGCCCATCGAGCTGCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1426
Db 4341 CCGCAAGTGGAGCGTGCAGCCCATCGTGTGCCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 4400

Tue Feb 11 14:53:47 2003

Qy	1427	ACATCAGAAAGCTGTGGCGAAAGCTGAAC	TGGCGGACGACCAAGTCTACCCCGCATCAAGG	1488
Db	4401	ACATCAGAAAGCTGTGGCGAAAGCTGAAC	TGGCGGACGACCAAGTCTACCCCGCATCAAGG	4460
Qy	1487	TGCGCGAGCTGTGCAGAGCTGTGCGCGGCG	CGAAGCCCTGACCGACATCGTGCCTCTGA	1546
Db	4461	TGCGCGAGCTGTGCAGAGCTGTGCGCGGCG	CGAAGCCCTGACCGAGGTGGTGCCTCTGA	4520
Qy	1547	CCGAGGAGCGCAGAGCTGGAGCTGGCGGAG	AACCGCGAGATCCTGCGCGAGCCCGTGCAG	1606
Db	4521	CCGAGGAGCGCAGAGCTGGAGCTGGCGGAG	AACCGCGAGATCCTGAGAGACCCGTCGACG	4580
Qy	1607	GCCTGTACTACGACCCCGACCAAGGACCT	TGGTGGCGGAGATCCAGAAGCAGGGCCAGACC	1666
Db	4581	GCCTGTACTACGACCCCGACCAAGGACCT	TGGTGGCGGAGATCCAGAAGCAGGGCCAGACC	4640
Qy	1667	AGTGGACCTACCAAGATCTACCAAGGAGCC	CTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	1726
Db	4641	AGTGGACCTACCAAGATCTACCAAGGAGCC	CTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	4700
Qy	1727	AGATGGCGCACCCCGCACCAACACGACGT	TGAAGCAGCTACCGAGCGCCGTGCAGAAATCG	1786
Db	4701	GCATGAAGGCGCCCGCACCAACACGACGT	TGAAGCAGCTACCGAGCGCCGTGCAGAAATCG	4760
Qy	1787	CCATGGAGAGCATCTGTGTATCTGGGCGA	AGACCCCAAGCTTCGCGCTGCCCATCCAGAAAG	1846
Db	4761	CCACCGAGAGCATCTGTGTATCTGGGCGA	AGACCCCAAGCTTCGCGCTGCCCATCCAGAAAG	4820
Qy	1847	AGACCTGGGAGACCTGTGTGGACCGACT	ACTTGGCAGGCCACCTGGATCCCAGTGGGAGT	1906
Db	4821	AGACCTTGGGAGGCGCTGTGTGGACCGAGT	ACTTGGCAGGCCACCTGGATCCCAGTGGGAGT	4880
Qy	1907	TGCTGAACACCCCGCCCTGTGTGAAGCT	TGTGTGTACCAAGCTTGGAGAAAGGCCCATCTCG	1966
Db	4881	TCGTGAACACCCCGCCCTGTGTGAAGCT	TGTGTGTACCAGCTTGGAGAAAGGCCCATCTCG	4940
Qy	1967	GGCCGAGACCTTCTACGTGTGACGGCGCG	CGCCCAACCGCGAGACCAAGATCGGCAAGCGCG	2026
Db	4941	GGCCGAGACCTTCTACGTGTGACGGCGCG	CGCCCAACCGCGAGACCAAGATCGGCAAGCGCG	5000
Qy	2027	GCTACTGTACCGACCGGGCGGCAAGAGAT	GTGTAGCGCTGACCGAGACCAACCAACCA	2086
Db	5001	GCTACTGTACCGACCGGGCGGCGCCCAAG	AGTGTGTGCCCTTGACCGACCAACCAACCA	5060
Qy	2087	AGACCGAGCTGCAGGCCATCCAGCTTGGC	CCCTCTGAGGACAGCGGCGAGGTGAAACATCG	2146
Db	5061	AGACCGAGCTGCAGGCCATCCAGCTTGGC	CCCTCTGAGGACAGCGGCGGTGAAACATCG	5120
Qy	2147	TGACCGACAGCCAGTACGCCCTTGGGCAT	CTATC2AGGCCAGCCGACAGAGCGAGCG	2206
Db	5121	TGACCGACAGCCAGTACGCCCTTGGGCAT	CTATC2AGGCCAGCCGACAGAGCGAGCG	5180
Qy	2207	AGCTGGTGAACAGATCATCGAGCAGCTGT	CAAGAAGGAGAGGTGTACCTGAGTGGG	2266
Db	5181	AGCTGGTGAACAGATCATCGAGCAGCTGT	CAAGAAGGAGAGGTGTACCTGAGTGGG	5240
Qy	2267	TGCCCCGCCCAAGGGCATCGCGCGCAAC	AGCAGCAGATCTGACAGCTGGTGTGAGCAAGGGCA	2326
Db	5241	TGCCCCGCCCAAGGGCATCGCGCGCAAC	AGCAGCAGATCTGACAGCTGGTGTGAGCAAGGGCA	5300
Qy	2327	TCCGCAAGGTGCTGTCTCTGTGACGGCAT	CTGATGGCGGATCTGTGATCTACCACTA	2381
Db	5301	TCCGCAAGGTGCTGTCTCTGTGACGGCAT	CTGATGGCGGATCTGACAGAGGCCCAAGGAGGACGACGAGAA	5355

Search completed: February 10, 2003, 17:59:44
Job time : 6356.78 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 10:51:03 ; Search time 489.804 Seconds

(without alignments)
11296.684 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcgagccaccatggccga.....gggtgacacgggtgaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	2457	24	ABL39961
2	2436.2	99.2	2463	24	ABL39960
3	2415.4	98.3	2469	24	ABL39959
4	2040	83.0	2300	21	AAA70480
5	2019.2	82.2	2306	21	AAA70479
6	2000.8	81.4	9166	24	ABK91616
7	1998.4	81.3	2312	21	ABK91622
8	1992.4	81.1	9788	24	ABK91622
9	1978.4	80.5	9169	24	ABK91617

10	1975	80.4	9194	24	ABK91612	Modified HIV prote
11	1973.4	80.3	9194	24	ABK91611	Modified HIV prote
12	1973.4	80.3	12411	24	ABK91613	Modified HIV prote
13	1970	80.2	9785	24	ABK91624	Modified HIV prote
14	1966.6	80.0	9167	24	ABK91619	Modified HIV prote
15	1966.6	80.0	9170	24	ABK91614	Modified HIV prote
16	1966.6	80.0	9782	24	ABK91621	Modified HIV prote
17	1966.6	80.0	9783	24	ABK91620	Modified HIV prote
18	1966.6	80.0	9792	24	ABK91618	Modified HIV prote
19	1965	80.0	9189	24	ABK91607	Modified HIV prote
20	1945.4	79.2	3015	24	ABL39983	Synthetic Pol poly
21	1928.8	78.5	3009	24	ABL40024	HIV Gag-polymerase
22	1915.8	78.0	4319	21	AAA70414	Synthetic Pol poly
23	1882.2	76.6	8908	21	RAZ52055	Packaging construc
24	1878.6	76.5	1965	24	ABL39989	Synthetic protease
25	1872.8	76.2	2299	21	AAA70478	HIV FS(+)-Proteinac
26	1859.8	75.7	2577	24	RAA141592	HIV-1 subtype C is
27	1852.2	75.4	4343	24	RAA167896	Nucleotide sequenc
28	1852	75.4	2305	21	AAA70477	HIV FS(+)-Proteinac
29	1848.2	75.2	4341	24	RAA167897	Nucleotide sequenc
30	1827.6	74.4	1978	24	ABL39991	Synthetic protease
31	1825.6	74.3	7897	24	ABK91609	Modified HIV prote
32	1762	71.7	3012	21	RAZ52054	Codon optimised Hu
33	1757.2	71.5	9407	24	ABK91623	Modified HIV prote
34	1607.4	65.4	4307	20	AAZ08740	HIV gagpol-SVNgp C
35	1607.4	65.4	4307	21	AAA93972	Human immunodefici
36	1607.4	65.4	4307	22	AAC86876	Nucleotide sequenc
37	1599.4	65.1	4307	22	AAH43696	Codon optimised HI
38	1599.4	65.1	4327	21	AAA93984	HIV partial leader
39	1599.4	65.1	4353	21	AAA93983	HIV partial leader
40	1599.4	65.1	4642	21	AAA93982	HIV complete leade
41	1599.4	65.1	9772	22	AAH43699	Codon optimised ga
42	1596	65.0	1668	24	ABL39999	Synthetic RT polyn
43	1554.4	63.3	1680	24	ABL39997	Synthetic RT polyn
44	1539.2	62.6	4338	22	AAH22806	DNA sequence of a
45	1539.2	62.6	8366	22	AAH22810	DNA sequence of a

ALIGNMENTS

RESULT 1

ABL39961
ID ABL39961 standard; DNA; 2457 BP.
XX
AC ABL39961;
XX
DT 15-MAY-2002 (first entry)
XX
DE Synthetic construct PR975YMMW SEQ ID NO:32.
XX
KW Human immunodeficiency virus type C; antigenic HIV type C protein; immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef; immunostimulant; gene therapy; gene; ds.
XX
OS Human immunodeficiency virus type C.
OS Synthetic.
XX
FN WO200204493-A2.
XX
PD 17-JAN-2002.
XX
PF 05-JUL-2001; 2001WO-US21241.
XX
PR 05-JUL-2000; 2000US-0610313.
XX
PA (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
PI Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX
DR WPI; 2002-154920/20.
XX

PT New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -

PS Claim 1; Fig 10; 233pp; English.

XX The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (I). (I) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.

XX Sequence 2457 BP; 566 A; 837 C; 754 G; 300 T; 0 other;

SQ Query Match 100.0%; Score 2457; DB 24; Length 2457;

Best Local Similarity 100.0%; Pred. No. 1.9e-297;

Matches 2457; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGCGCCACCATGGCGGAGGCGCATGAGCCAGGCGCCACAGCCGCAACATCTCTGATGCAG 60
DB 1 GTGAGCGCCACCATGGCGGAGGCGCATGAGCCAGGCGCCACAGCCGCAACATCTCTGATGCAG 60
QY 61 CGCAGCAACTTCAAGGGCCCCAGCGCATCATCAAGTGTTCACCTGCGCGCAAGGAGGC 120
DB 61 CGCAGCAACTTCAAGGGCCCCAGCGCATCATCAAGTGTTCACCTGCGCGCAAGGAGGC 120
QY 121 CACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAG 180
DB 121 CACATCGCCCGCAACTGCGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAG 180
QY 181 GGGCCACAGATGAAGGACTGCACCGAGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
DB 181 GGGCCACAGATGAAGGACTGCACCGAGCGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
QY 241 TTTCCCGCAGGGCAAGGCCGCGAGTTTCCCGCAGGAGCAGCAACCGCGCAACACGCCACC 300
DB 241 TTTCCCGCAGGGCAAGGCCGCGAGTTTCCCGCAGGAGCAGCAACCGCGCAACACGCCACC 300
QY 301 AGCGCGAGCTGCAAGTGCAGCGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCG 360
DB 301 AGCGCGAGCTGCAAGTGCAGCGCGCAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCG 360
QY 361 GGCACCTGAACCTTCCCGCAGATCAACCTGTGGCAGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 GGCACCTGAACCTTCCCGCAGATCAACCTGTGGCAGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 GCGCGCCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 GCGCGCCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 ATGAGCTGCGCGCGCAAGTGGAGCGCAAGATGATGCGCGCGCATCGCGCGCTTCATCAAG 540
DB 481 ATGAGCTGCGCGCGCAAGTGGAGCGCAAGATGATGCGCGCGCATCGCGCGCTTCATCAAG 540
QY 541 GTGCGCGCAGTACGACCGCATTCCTGTAGATCTGCGGCAAGAGGCGCATCGGACCCGTG 600
DB 541 GTGCGCGCAGTACGACCGCATTCCTGTAGATCTGCGGCAAGAGGCGCATCGGACCCGTG 600
QY 601 CTGATCGGCG 660
DB 601 CTGATCGGCG 660
QY 661 ACCCTGAACCTTCCCGCAGTACGACCGCATCGAGCGCGTGGCGGTGAAGCTGAAGCGCGG 720
DB 661 ACCCTGAACCTTCCCGCAGTACGACCGCATCGAGCGCGTGGCGGTGAAGCTGAAGCGCGG 720
QY 721 GACGGCG 780

DB 721 GACGGCG 780
QY 781 ATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAGATCGGCCCGCGCGAGAGCCCTAC 840
DB 781 ATCTGCGAGGAGATGGAGAGGAGGCAAGATCACCAGATCGGCCCGCGCGAGAGCCCTAC 840
QY 841 AACACCCCGCTGTTCGCCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
DB 841 AACACCCCGCTGTTCGCCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 TTCCGCGAGCTGAACAG 960
DB 901 TTCCGCGAGCTGAACAG 960
QY 961 CCGCGCGCGCTGAAG 1020
DB 961 CCGCGCGCGCTGAAG 1020
QY 1021 AGCGTGCCTGCGAGGAG 1080
DB 1021 AGCGTGCCTGCGAGGAG 1080
QY 1081 AAG 1140
DB 1081 AAG 1140
QY 1141 CCGAG 1200
DB 1141 CCGAG 1200
QY 1201 GAGATCGTGTATACAG 1260
DB 1201 GAGATCGTGTATACAG 1260
QY 1261 CCGCGCGAGATCGAGGAG 1320
DB 1261 CCGCGCGAGATCGAGGAG 1320
QY 1321 AAG 1380
DB 1321 AAG 1380
QY 1381 GTGAGCG 1440
DB 1381 GTGAGCG 1440
QY 1441 GTGGGCAAGCTGAACCTGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB 1441 GTGGGCAAGCTGAACCTGCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 AAGCTGTGCGAG 1560
DB 1501 AAGCTGTGCGAG 1560
QY 1561 CTGAGCTGCGAG 1620
DB 1561 CTGAGCTGCGAG 1620
QY 1621 CCGAG 1680
DB 1621 CCGAG 1680
QY 1681 ATCTTACAG 1740
DB 1681 ATCTTACAG 1740
QY 1741 CACACCAACAG 1800
DB 1741 CACACCAACAG 1800
QY 1801 GTGATCTGGGCGAG 1860

Db 1801 GTGATCTGGGGCAAGACCCCCAAGTTCGCCCTGCCATCCAGAGGAGACCTGGGAGACC 1860
Qy 1861 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCC 1920
Db 1861 TGGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGAGTTCGTGAACACCCCC 1920
Qy 1921 CCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCGCGGAGACCTTC 1980
Db 1921 CCCCTGGTGAAGCTGTGTACAGCTGGAGAGAGGCCCATCATCGGCGCGGAGACCTTC 1980
Qy 1981 TACGTGGAGCGCGCCCAACCGCAGACCAAGATCGCAAGCGCGGTACGTGACCGAC 2040
Db 1981 TACGTGGAGCGCGCGCCCAACCGCAGACCAAGATCGCAAGCGCGGTACGTGACCGAC 2040
Qy 2041 CGGGCCGCGCAGAAAGATCGTAGCCTGACCGAGACCAACCAAGAACCGAGCTGCAG 2100
Db 2041 CGGGCCGCGCAGAAAGATCGTAGCCTGACCGAGACCAACCAAGAACCGAGCTGCAG 2100
Qy 2101 GCCATCCAGCTGGCCCTGCAGACAGCGGACGAGGTGAACATCGTGACCGACACCGCAG 2160
Db 2101 GCCATCCAGCTGGCCCTGCAGACAGCGGACGAGGTGAACATCGTGACCGACACCGCAG 2160
Qy 2161 TACGCCCTGGGCATCATCCAGGCCAGCCGACACAGAGCGAGCGAGCTGTGTGAACCGAG 2220
Db 2161 TACGCCCTGGGCATCATCCAGGCCAGCCGACACAGAGCGAGCGAGCTGTGTGAACCGAG 2220
Qy 2221 ATCATCGAGCAGCTGATCAAGAAGAGAGAGTGTACCTGAGCTGGTGGCGCCCAACAAG 2280
Db 2221 ATCATCGAGCAGCTGATCAAGAAGAGAGAGTGTACCTGAGCTGGTGGCGCCCAACAAG 2280
Qy 2281 GGCATCGCGGCGAAGCAGATGCACAAAGCTGTGTGAGCAAGGGGCATCCGCAAGGTGCTG 2340
Db 2281 GGCATCGCGGCGAAGCAGATGCACAAAGCTGTGTGAGCAAGGGGCATCCGCAAGGTGCTG 2340
Qy 2341 TTCTTGGAGCGCATCGATGGCGGCATCGTATCTACCACTACATGAGCAGACCTGTACCTG 2400
Db 2341 TTCTTGGAGCGCATCGATGGCGGCATCGTATCTACCACTACATGAGCAGACCTGTACCTG 2400
Qy 2401 GGCAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457
Db 2401 GGCAGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGGCTAGCACCGGTGAATTC 2457

RESULT 2
ABL39960
ID ABL39960 standard; DNA: 2463 BP.
AC ABL39960;
XX
AC
XX
DT 15-MAY-2002 (first entry)
XX
XX Synthetic construct PR975YM SEQ ID NO:31.
DE Human immunodeficiency virus type C; antigenic HIV type C protein; immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef; immunostimulant; gene therapy; gene; ds.
KW
KW
XX Human immunodeficiency virus type C.
OS Synthetic.
OS
XX WO200204493-A2.
XX
XX 17-JAN-2002.
XX
XX 05-JUL-2001; 2001WO-US21241.
XX
XX 05-JUL-2000; 2000US-0610313.
XX
XX (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX
XX Zur Wegede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
PI
XX

DR
XX
PT
PT
XX
PS
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

WPI; 2002-154920/20.

New polynucleotides encoding antigenic HIV Type C polypeptides, useful in applications including DNA immunization or generation of packaging cell lines, particularly in gene therapy

Claim 1; Fig 9; 233pp; English.

The present invention describes expression cassettes comprising a polynucleotide sequence encoding a polypeptide comprising immunogenic HIV type C polypeptides. The expression cassettes comprise any of the HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or Nef (1). (I) have immunostimulant activity and can be used in gene therapy. The HIV type C polynucleotides are useful in applications including DNA immunisation, generation of packaging cell lines, and production of HIV type C proteins. The polynucleotides are particularly useful in gene therapy and DNA immunisation applications. ABL39942 to ABL40054 and ABL06204 to ABL06215 represent sequences used in the exemplification of the present invention.

Sequence 2463 BP; 567 A; 835 C; 759 G; 302 T; 0 other;

Query Match 99.2%; Score 2436.2; DB 24; Length 2463;
Best Local Similarity 99.6%; Pred. No. 7.2e-295;
Matches 2454; Conservative 0; Mismatches 3; Indels 6; Gaps 1;

Qy 1 GTCGAGCGCCACCATCGCGAGGCCATGAGCCAGGCCACCGCCCAACATCTGTATGCAG 60
Db 1 GTCGAGCGCCACCATCGCGAGGCCATGAGCCAGGCCACCGCCCAACATCTGTATGCAG 60
Qy 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGCG 120
Db 61 CGCAGCAACTTCAAGGGCCCCAAGCGCATCATCAAGTCTTCAACTCGCGCAAGGAGGCG 120
Qy 121 CACATCGCGCGCACTGCGCGCCCGCCCGCAAGAGGGCTGTGGAAGTGCAGGCAAGGAG 180
Db 121 CACATCGCGCGCACTGCGCGCCCGCCCGCAAGAGGGCTGTGGAAGTGCAGGCAAGGAG 180
Qy 181 GSCCACCAGATGAAGAGCTGACCGCGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
Db 181 GSCCACCAGATGAAGAGCTGACCGCGCGCCAGCGCAACTTCTTCGCGAGGACCTGGCC 240
Qy 241 TTCCCCCAGGCGAAGCGCCCGGAGTTCCCGAGCAGAGAGCGCGCCCAACAGCCCGCCAC 300
Db 241 TTCCCCCAGGCGAAGCGCCCGGAGTTCCCGAGCAGAGAGCGCGCCCAACAGCCCGCCAC 300
Qy 301 AGCCCGAGCTGCGAGTGGCGGGCGACAAACCCCGCAGGAGCGCGCGCGAGCGCCAG 360
Db 301 AGCCCGAGCTGCGAGTGGCGGGCGACAAACCCCGCAGGAGCGCGCGCGAGCGCCAG 360
Qy 361 GGCACCTGAACTTCCCGCAGATCACCTGTGCGAGCGCCCGCTGTGAGCATCAAGGTG 420
Db 361 GGCACCTGAACTTCCCGCAGATCACCTGTGCGAGCGCCCGCTGTGAGCATCAAGGTG 420
Qy 421 GCGGCCAGATCAAGAGGCGCTGTGAGACACCGCGCGCGAGACACCGTGTGGAGGAG 480
Db 421 GCGGCCAGATCAAGAGGCGCTGTGAGACACCGCGCGCGAGACACCGTGTGGAGGAG 480
Qy 481 ATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATCGCGGGCATCGCGGGCTTCATCAAG 540
Db 481 ATGAGCTTCCCGGCAAGTGAAGCCCAAGATGATCGCGGGCATCGCGGGCTTCATCAAG 540
Qy 541 GTGGCCAGTACGACCATCTGATCGAGATCTCGGCAAGAGGCCATCGGCCCGCTG 600
Db 541 GTGGCCAGTACGACCATCTGATCGAGATCTCGGCAAGAGGCCATCGGCCCGCTG 600
Qy 601 CTGATCGGCCCGACCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGGCTGC 660
Db 601 CTGATCGGCCCGACCCCGTGAACATCATCGGCCCGCAACATGCTGACCCAGCTGGGCTGC 660
Qy 661 ACCCTGAAGTTCGCCCATCAGCCCCATCGAGACCGCTGCCGTGAAGCTGAAGCCCGCATG 720
Db 661 ACCCTGAAGTTCGCCCATCAGCCCCATCGAGACCGCTGCCGTGAAGCTGAAGCCCGCATG 720

```
QY 721 GACGGCCCCAAGGTGAAGCAGTGGCCCTCACCAGGAGAGATCAAGGCCCTGACCGCC 780
Db 721 GACGGCCCCAAGGTGAAGCAGTGGCCCTCACCAGGAGAGATCAAGGCCCTGACCGCC 780
QY 781 ATCTGGAGAGATGGAGAAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCTAC 840
Db 781 ATCTGGAGAGATGGAGAAGGAGGCAAGATCACCAGATCGGCCCGGAGAACCCCTAC 840
QY 841 AACACCCCGCTGTTCGCCATCAAGAGAAGAGCAGCACCAGTGGCGCAAGCTGGTGGAC 900
Db 841 AACACCCCGCTGTTCGCCATCAAGAGAAGAGCAGCACCAGTGGCGCAAGCTGGTGGAC 900
QY 901 TTCGGGAGCTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCGCCAC 960
Db 901 TTCGGGAGCTGAACAGCCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCGCCAC 960
QY 961 CCGCGCGCTGAAGAGAAGAGAGCGTGACCGTGTGGACGTGGCGCAGCGCTACTTC 1020
Db 961 CCGCGCGCTGAAGAGAAGAGAGCGTGACCGTGTGGACGTGGCGCAGCGCTACTTC 1020
QY 1021 AGCGTGCCCTGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCAAC 1080
Db 1021 AGCGTGCCCTGACGAGGACTTCCGCAAGTACACCGCTTACCATCCCGCAGCATCAAC 1080
QY 1081 AACGAGACCCCGGCATCCGCTACCAAGTACAACGTGCTGCCCGAGGCTGGAAGGGCAGC 1140
Db 1081 AACGAGACCCCGGCATCCGCTACCAAGTACAACGTGCTGCCCGAGGCTGGAAGGGCAGC 1140
QY 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC 1200
Db 1141 CCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCCGCAACCCC 1200
QY 1201 GAGTGTGATCTACAGGCCCCCTGTAGCTGGGAGCGACCTTGGAGATCGGCGCAGC 1260
Db 1201 GAGTGTGATCTACAGGCCCCCTGTAGCTGGGAGCGACCTTGGAGATCGGCGCAGC 1260
QY 1261 CGCGCCAAGATCGAGAGCTGCGCAGCAGCCTGCTGGCTGGGGCTTACACACCCCGCAG 1320
Db 1261 CGCGCCAAGATCGAGAGCTGCGCAGCAGCCTGCTGGCTGGGGCTTACACACCCCGCAG 1320
QY 1321 AAGAAGCAGAGAGAGCCCGCTTCTGCCAT-----CGAGCTGACCCCGCAGAAC 1374
Db 1321 AAGAAGCAGAGAGAGCCCGCTTCTGCCAT-----CGAGCTGACCCCGCAGAAC 1374
QY 1375 TGGACCGTCAGCCCATCGAGCTGCCCGAGAGAGAGCTGGACCTGACGACATCCAG 1434
Db 1375 TGGACCGTCAGCCCATCGAGCTGCCCGAGAGAGAGCTGGACCTGACGACATCCAG 1434
QY 1435 AAGCTGTGGGCAAGCTGAAGTGGGCGCAGCAGATCTACCCCGCATCAAGGTGCGCGAG 1494
Db 1435 AAGCTGTGGGCAAGCTGAAGTGGGCGCAGCAGATCTACCCCGCATCAAGGTGCGCGAG 1494
QY 1495 CTGTGAAGCTGTGCGCGCGCAGAGGCGCTGACCGACATCTGCCCCCTGACCGAGGAG 1554
Db 1495 CTGTGAAGCTGTGCGCGCGCAGAGGCGCTGACCGACATCTGCCCCCTGACCGAGGAG 1554
QY 1501 CTGTGAAGCTGTGCGCGCGCAGAGGCGCTGACCGACATCTGCCCCCTGACCGAGGAG 1560
Db 1501 CTGTGAAGCTGTGCGCGCGCAGAGGCGCTGACCGACATCTGCCCCCTGACCGAGGAG 1560
QY 1555 GCGAGCTGAGCTGGCGGAGACCGGAGATCTGCGCGAGCCCTGCGACGGCTGTAC 1614
Db 1555 GCGAGCTGAGCTGGCGGAGACCGGAGATCTGCGCGAGCCCTGCGACGGCTGTAC 1614
QY 1561 GCGAGCTGAGCTGGCGGAGACCGGAGATCTGCGCGAGCCCTGCGACGGCTGTAC 1620
Db 1561 GCGAGCTGAGCTGGCGGAGACCGGAGATCTGCGCGAGCCCTGCGACGGCTGTAC 1620
QY 1615 TAGGACCCCAAGGAGCCTTGGTGGCGGAGATCCAGAAGCAGGCGCCACGACCTGTGACC 1674
Db 1615 TAGGACCCCAAGGAGCCTTGGTGGCGGAGATCCAGAAGCAGGCGCCACGACCTGTGACC 1674
QY 1675 TACAGATCTACAGGAGCCCTTCAAGAACTGAAGACCGGCAAGTACGCCAAGATCGCG 1734
Db 1675 TACAGATCTACAGGAGCCCTTCAAGAACTGAAGACCGGCAAGTACGCCAAGATCGCG 1734
QY 1735 ACCGCCACACCAAGCAGCTGAAGCAGCTGACGAGCGCTGAGAGAGATCGCCATCGAG 1794
Db 1735 ACCGCCACACCAAGCAGCTGAAGCAGCTGACGAGCGCTGAGAGAGATCGCCATCGAG 1794
QY 1741 ACCGCCACACCAAGCAGCTGAAGCAGCTGACGAGCGCTGAGAGAGATCGCCATCGAG 1800
Db 1741 ACCGCCACACCAAGCAGCTGAAGCAGCTGACGAGCGCTGAGAGAGATCGCCATCGAG 1800
```

```
QY 1795 AGCATGTGATCTGGGCAAGACCCCAAGTTCGGCTGCCCATCCAGAAGGAGACCTGG 1854
Db 1801 AGCATGTGATCTGGGCAAGACCCCAAGTTCGGCTGCCCATCCAGAAGGAGACCTGG 1860
QY 1855 GAGACCTGTGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1914
Db 1861 GAGACCTGTGTGGACCGACTACTGGCAGGCCACCTGGATCCCGAGTGGGAGTTCGTGAAC 1920
QY 1915 ACCCCGCCCTGTGTAAGCTGTGTACCACTGGAGAGAGCCCATCATCGGCGCGAG 1974
Db 1921 ACCCCGCCCTGTGTAAGCTGTGTACCACTGGAGAGAGCCCATCATCGGCGCGAG 1980
QY 1975 ACCTTCTACCTGGAGCGCGCCCAACCGCAGACCAAGATCGCAAGCGCGGTACCTG 2034
Db 1981 ACCTTCTACCTGGAGCGCGCCCAACCGCAGACCAAGATCGCAAGCGCGGTACCTG 2040
QY 2035 ACCGACCGGGCGCGCAGAGAAGATCGTGAGCTGACCGAGACCAACCAAGAGACCGAG 2094
Db 2041 ACCGACCGGGCGCGCAGAGAAGATCGTGAGCTGACCGAGACCAACCAAGAGACCGAG 2100
QY 2095 CTGACGCCATCCAGCTGGCCCTGACGAGCAGCGGCGAGCGGTGAACATCGTGACCGAC 2154
Db 2101 CTGACGCCATCCAGCTGGCCCTGACGAGCAGCGGCGAGCGGTGAACATCGTGACCGAC 2160
QY 2155 AGCAGTACCCCTGGGCATCATCCAGCGCCAGCCGACAGCGAGCGAGCTGGTG 2214
Db 2161 AGCAGTACCCCTGGGCATCATCCAGCGCCAGCCGACAGCGAGCGAGCTGGTG 2220
QY 2215 AACGATCATTCAGCAGCTGATCAAGAAGAGAGTGTACCTGAGCTGGTGCCCGCC 2274
Db 2221 AACGATCATTCAGCAGCTGATCAAGAAGAGAGTGTACCTGAGCTGGTGCCCGCC 2280
QY 2275 CACAAGGCGATCGCGGCAAGCAGCAGATCGACAAGCTGGTGAGCAAGGCGCATCCGCAAG 2334
Db 2281 CACAAGGCGATCGCGGCAAGCAGCAGATCGACAAGCTGGTGAGCAAGGCGCATCCGCAAG 2340
QY 2335 GTGTGTTCCTGGACGCGCATCGATGCGCGCATCTGATCTACCAAGTACATGGAGACCTG 2394
Db 2341 GTGTGTTCCTGGAGGCGCATCGATGCGCGCATCTGATCTACCAAGTACATGGAGACCTG 2400
QY 2395 TAGTGGGAGCGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2454
Db 2401 TAGTGGGAGCGCGCGCCCTAGGATCGATTAAAGCTTCCCGGGCTAGCACCGGTGAA 2460
QY 2455 TTC 2457
Db 2461 TTC 2463

RESULT 3
ABL39959
ID ABL39959 standard; DNA; 2469 BP.
XX ABL39959;
XX AC
XX DT
XX 15-MAY-2002 (first entry)
XX Synthetic construct PR975(+) SEQ ID NO:30.
XX Human immunodeficiency virus type C; antigenic HIV type C protein;
KW immunogenic; immunisation; gag; pol; vif; vpr; tat; rev; vpu; env; nef;
KW immunostimulant; gene therapy; gene; ds.
XX Human immunodeficiency virus type C.
OS Synthetic.
XX WO200204493-A2.
XX 17-JAN-2002.
XX 05-JUL-2001; 2001WO-US21241.
XX 05-JUL-2000; 2000US-0610313.
PR
```

XX (CHIR) CHIRON CORP.
PA (UYST-) UNIV STELLENBOSCH.
XX Zur Megede J, Barnett SW, Engelbrecht S, Van Rensburg EJ;
XX WPI; 2002-154920/20.
XX New polynucleotides encoding antigenic HIV Type C polypeptides, useful
PT in applications including DNA immunization or generation of packaging
PT cell lines, particularly in gene therapy -
XX Claim 1; Fig 8; 233pp; English.
XX The present invention describes expression cassettes comprising a
CC polynucleotide sequence encoding a polypeptide comprising immunogenic
CC HIV type C polypeptides. The expression cassettes comprise any of the
CC HIV type C sequences encoding Gag, Pol, Vif, Vpr, Tat, Rev, Vpu, Env or
CC Nef (1). (1) have immunostimulant activity and can be used in gene
CC therapy. The HIV type C polynucleotides are useful in applications
CC including DNA immunisation, generation of packaging cell lines, and
CC production of HIV Type C proteins. The polynucleotides are particularly
CC useful in gene therapy and DNA immunisation applications. ABL39942 to
CC ABL40054 and ABL06204 to ABL06215 represent sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 2469 BP; 571 A; 833 C; 761 G; 304 T; 0 other;
Query Match 98.3%; Score 2415.4; DB 24; Length 2469;
Best Local Similarity 99.3%; Pred. No. 2.8e-292;
Matches 2451; Conservative 0; Mismatches 5; Indels 12; Gaps 2;
QY 1 GTCGACGCCACATGGCGGAGGCCATGAGCCAGGCCACAGCGCCCAACATCCTGATGCAG 60
DB 1 GTCGACGCCACATGGCGGAGGCCATGAGCCAGGCCACAGCGCCCAACATCCTGATGCAG 60
QY 61 GCGAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGGC 120
DB 61 GCGAGCAACTTCAAGGCCCCCAAGCGCATCATCAAGTCTTCAACTGCGGCAAGAGGGC 120
QY 121 CACATCGCCGCAACTGCGGCGCCCGCCGCAAGAGGGCTGCTGAAGTGGCGCAAGAG 180
DB 121 CACATCGCCGCAACTGCGGCGCCCGCCGCAAGAGGGCTGCTGAAGTGGCGCAAGAG 180
QY 181 GGCCACCATGTAAGGACTGCAAGCGGCGGCGGCAAGTCTTCCGCGAGGACCTTGCC 240
DB 181 GGCCACCATGTAAGGACTGCAAGCGGCGGCGGCAAGTCTTCCGCGAGGACCTTGCC 240
QY 241 TTCCCGGAGGAGGCGCGGAGTTCGCCAGCGAGCAACCGCGGCAAGCGCCCAAC 300
DB 241 TTCCCGGAGGAGGCGCGGAGTTCGCCAGCGAGCAACCGCGGCAAGCGCCCAAC 300
QY 301 AGCGCGAGCTGCAAGTGGCGGCGCAACCCCGAGCGAGGCGCGGCGGCGGCGGCG 360
DB 301 AGCGCGAGCTGCAAGTGGCGGCGCAACCCCGAGCGAGGCGCGGCGGCGGCGGCG 360
QY 361 GGCACTTGAACCTTCCCGGAGATCACCTGTGGCAGCGCCCGCTTGGTGAAGTGA 420
DB 361 GGCACTTGAACCTTCCCGGAGATCACCTGTGGCAGCGCCCGCTTGGTGAAGTGA 420
QY 421 GGCGGCGAGATCAAGGAGGCGCTGTCGACACCGCGGCGGCGGCGGCGGCGGCG 480
DB 421 GGCGGCGAGATCAAGGAGGCGCTGTCGACACCGCGGCGGCGGCGGCGGCGGCG 480
QY 481 ATGAGCTGCGCGCAAGTGGAGGCGCAAGATGATCGCGGCGGCGGCGGCGGCGG 540
DB 481 ATGAGCTGCGCGCAAGTGGAGGCGCAAGATGATCGCGGCGGCGGCGGCGGCGG 540
QY 541 GTGCGGAGTACGACCATGATCTGATCGAGATCTGCGGCGGCAAGAGGCGGCGGCG 600
DB 541 GTGCGGAGTACGACCATGATCTGATCGAGATCTGCGGCGGCAAGAGGCGGCGGCG 600
QY 601 CTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660

DB 601 CTGATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 661 ACCCTGAACCTTCCCGATCAGCCCATCAGACCGCTGCGGCGGCAAGTGAAGCGCGGCG 720
DB 661 ACCCTGAACCTTCCCGATCAGCCCATCAGACCGCTGCGGCGGCAAGTGAAGCGCGGCG 720
QY 721 GACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
DB 721 GACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 780
QY 781 ATCTCGGAGGATGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 ATCTCGGAGGATGGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 841 AACACCCCGGCTTCCCGATCAGCCCATCAGAGAGGAGGCGGCGGCGGCGGCGGCGGCG 900
DB 841 AACACCCCGGCTTCCCGATCAGCCCATCAGAGAGGAGGCGGCGGCGGCGGCGGCGGCG 900
QY 901 TTCCGCGAGCTGGAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 901 TTCCGCGAGCTGGAACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 961 CCGCGCGGCGCTGGAAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 961 CCGCGCGGCGCTGGAAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
QY 1021 AGCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
DB 1021 AGCGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
QY 1081 AACGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140
DB 1081 AACGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1140
QY 1141 CCCAGCATCTTCCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 CCCAGCATCTTCCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 GAGATCGTGATCTACCA-----GGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1254
DB 1201 GAGATCGTGATCTACCAAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
QY 1255 CAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1314
DB 1261 CAGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
QY 1315 CCCGACAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1368
DB 1321 CCCGACAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
QY 1369 GACAGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1428
DB 1381 GACAGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
QY 1429 ATCCAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1488
DB 1441 ATCCAGAGCTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
QY 1489 GCGCAGCTGTGCAAGTGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1548
DB 1501 GCGCAGCTGTGCAAGTGTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
QY 1549 GAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1608
DB 1561 GAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
QY 1609 GTGTACTACGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1668
DB 1621 GTGTACTACGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1680
QY 1669 TGGACCTACGAGATCTACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1728

Db 541 CCGTGAAGCTGAAGCGGGGATGAGACGGCCCCAAGGTCAAGCACTGGCCCCCTGACCCGAGG 600
Qy 758 AGAAGATCAAGGCCCTGTACCGCCATCTCCGAGGAGATGGAAGAGGCGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTGTGTGAGATCTCCAGGAGATGGAGAAGGAGGCGAAGATCACCA 660
Qy 818 AGATCGGCCCGGAGAACCCCTTACAAACACCCCGGTGTTCCGCCATCAAGAAGAAGACAGCA 877
Db 661 AGATCGGCCCGGAGAACCCCTTACAAACACCCCGGTGTTCCGCCATCAAGAAGAAGACAGCA 720
Qy 878 CCAAGTGGCGCAAGCTGTGTGACATCTCCGAGCTGGAACAAGCGACCCAGACATCTCTGGG 937
Db 721 CCAAGTGGCGCAAGCTGTGTGACATCTCCGAGCTGGAACAAGCGACCCAGACATCTCTGGG 780
Qy 938 AGGTGCAGCTGGGATCCCCACCCCGCGCTGGAACAAGAAGAGCGTGAACCGTGC 997
Db 781 AGGTGCAGCTGGGATCCCCACCCCGCGCTGGAACAAGAAGAGCGTGAACCGTGC 840
Qy 998 TGGAGCTGGGCGAGCCCTTACTTCAAGCTGCGCCCTGGACAAGGACTTCCGCAAGTACACCG 1057
Db 841 TGGAGCTGGGCGAGCCCTTACTTCAAGCTGCGCCCTGGACAAGGACTTCCGCAAGTACACCG 900
Qy 1058 CCTTCACCATCCCGAGCATCAACACGAGACCCCGGCGATCCGCTACCAAGTACAACTGC 1117
Db 901 CCTTCACCATCCCGAGCATCAACACGAGACCCCGGCGATCCGCTACCAAGTACAACTGC 960
Qy 1118 TGCOCAGGGCTGGAAGGCGAGCCCGACATCTTCCAGAGCAGCATGACCAAGATCTCTGG 1177
Db 961 TGCOCAGGGCTGGAAGGCGAGCCCGACATCTTCCAGAGCAGCATGACCAAGATCTCTGG 1020
Qy 1178 AGCCCTTCGCGCGCGCAACCCCGAGATCGTGATCTTACCAGGCCCGCTGTACCTGGSCA 1237
Db 1021 AGCCCTTCGCGCAAGCAAAACCCCGACATCGTGATCTTACCAGGCCCGCTGTACCTGGSCA 1080
Qy 1238 GCGACCTTGGAGATCGGCAGCAGCCCGCCAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1297
Db 1081 GCGACCTTGGAGATCGGCAGCAGCCCGCCAGATCGAGGAGCTGCGCAAGCACCTGCTGC 1140
Qy 1298 GCTGGGGCTTCAACACCCCGCACAAAGACACCAAGAGAGCCCGCTTCTGCGCCATCG 1357
Db 1141 GCTGGGGCTTCAACACCCCGCACAAAGACACCAAGAGAGCCCGCTTCTGCGCCATCG 1200
Qy 1358 AGCTGCACCCCGACAAGTGGACCTGACGCCATCGAGCCCTGCGGAGAGAGCTGGA 1417
Db 1201 AGCTGCACCCCGACAAGTGGACCTGACGCCATCGAGCCCTGCGGAGAGAGCTGGA 1260
Qy 1418 CCGTGAAGGATCCAGAAGCTGTGGCAAGCTGACTGGCGCAGCCAGATCTTACCCCG 1477
Db 1261 CCGTGAAGGATCCAGAAGCTGTGGCAAGCTGACTGGCGCAGCCAGATCTTACCCCG 1320
Qy 1478 GCATCAAGGTGCGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCCCTGACCCAGACATCG 1537
Db 1321 GCATCAAGGTGCGCCAGCTGTGCAAGCTGTGCGCGCGCCCAAGGCCCTGACCCAGATCG 1380
Qy 1538 TGCCCTTGACCGAGGCGGAGCTGGAGCTGGCCGCAACCGGAGATCTGCGCGAGC 1597
Db 1381 TGCCCTTGACCGAGGCGGAGCTGGAGCTGGCCGCAACCGGAGATCTGCGCGAGC 1440
Qy 1598 CCGTGCAGCGGCTGCTACTAGACCCCGACAGAGACCTGTTGGCGGAGATCCAGAGCAGG 1657
Db 1441 CCGTGCAGCGGCTGCTACTAGACCCCGACAGAGACCTGTTGGCGGAGATCCAGAGCAGG 1500
Qy 1658 GCCACGACAGTGGACCTTACAGATCTTACCAGAGCCCTTCAAGAACCTGAAGACCGGCA 1717
Db 1501 GCCAGGCCAGTGGACCTTACAGATCTTACCAGAGCCCTTCAAGAACCTGAAGACCGGCA 1560
Qy 1718 AGTACGCCAAGATCGGCACCGCCCGCACACCAACAGCTGAAGAGCTGACCGAGCCGTCG 1777
Db 1561 AGTACGCCCGCATCGCGGGCGCCCGCACACCAACAGCTGAAGCAGCTGACCGAGCCGTCG 1620
Qy 1778 AGAAGATCGGCATGGAGAGCATCTGTATCTGGGCGAAGACCCCAAGTTCGCGCTGCCCA 1837
Db 1621 AGAAGTGGACCGAGAGCATCTGTATCTGGGCGAAGATCCCAAGTTCAGGCTGCCCA 1680

Qy 1838 TCCAGAAGGAGACCTGGGAGACCTGTTGACCGCACTACTGGCAGGCCACCTGATCCCCG 1897
Db 1681 TCCAGAAGGAGACCTGGGAGGCTGTGTGATGGAGTACTTGGCAGGCCACCTGGATCCCCG 1740
Qy 1898 AGTGGAGTTCGTGAACACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGC 1957
Db 1741 AGTGGAGTTCGTGAACACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGC 1800
Qy 1958 CCATCATGGCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCGAGACCAAGATCG 2017
Db 1801 CCATCATGGCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCGAGACCAAGCTGG 1860
Qy 2018 GCAAGCGCGGTACGTGACCGACCGCGCGCAGAAAGATCTGAGCCCTGACCGGACCA 2077
Db 1861 GCAAGCGCGGTACGTGACCGACCGCGCGCAGAAAGTGTGAGCATTCGCCGACACCA 1920
Qy 2078 CCAACCAAGAAGCCAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGCAGCGAGG 2137
Db 1921 CCAACCAAGAAGCCAGCTGACGAGCCATCCAGCTGGCCCTGCAGGACAGCGCAGCGAGG 1980
Qy 2138 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCGCATCATCCAGGCCAGCCCGACAAG 2197
Db 1981 TGAACATCGTGACCGACAGCCAGTACGCCCTGGGCGCATCATCCAGGCCAGCCCGACAAG 2040
Qy 2198 GCGAGAGCAGCTGTGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACC 2257
Db 2041 GCGAGAGCAGCTGTGTGAACACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACC 2100
Qy 2258 TGAAGTGGTGGCCCGCCACAAAGGCGATCGCGCGCAAGCAGAGATCGACAAAGCTGGTGA 2317
Db 2101 TGGCTGGTGGCCCGCCACAAAGGCGATCGCGCGCAAGCAGAGCTGGACAAAGCTGGTGA 2160
Qy 2318 GCAAGGCGATCCGCAAGGTGTCTTCTGGAGCGCATCGATGGCGCATCTGATCTTACC 2377
Db 2161 GCGCGGCGATCCGCAAGGTGTCTTCTTGAAGCGCATCGATGGCGCATCTGATCTTACC 2220
Qy 2378 AGTACATGACGACCTGTACGTGGCGAGCGCGCCCTAGGATCGATTAAAGCTTCCCG 2437
Db 2221 AGTACATGACGACCTGTACGTGGCGAGCGCGCCCTAGGATCGATTAAAGCTTCCCG 2377
Qy 2438 GGGCTAGCACCGGTGAATTC 2457
Db 2281 GGGCTAGCACCGGTGAATTC 2300

RESULT 5

AAA70479

ID AAA70479 standard; DNA; 2306 BP.

AC AAA70479;

XX

XX 28-NOV-2000 (first entry)

DT

XX HIV FS(-)_ProtMod_rNpt_YM coding sequence.

DE

XX HIV-1; AIDS; Gag-reverse transcriptase; vaccine; expression cassette; ss.

KW

XX Human immunodeficiency virus type 1.

OS

XX Synthetic.

PN

XX WO200039302-A2.

XX

PD 06-JUL-2000.

XX

PF 30-DEC-1999; 99WO-US31245.

XX

PR 31-DEC-1998; 98US-0114495.

PR

PR 01-DEC-1999; 99US-0168471.

XX

PA (CHIR) CHIRON CORP.

XX

PI Barnett S, Zur Megede J, Srivastava I, Lian Y, Hartog K, Liu H;

PI Greer C, Selby M, Walker C;
XX WPI; 2000-452400/39.
XX
XX Expression cassettes encoding the human immunodeficiency virus (HIV)
PT Gag-containing polypeptide useful for vaccinating against HIV
PT Infections and acquired immunodeficiency syndrome (AIDS) -
XX
XX Claim 7; Fig 73; 39lpp; English.
PS
PS The present sequence is the coding sequence of a synthetic HIV
XX
CC Gag-reverse transcriptase expression cassette, FS(-).ProtMod.Rtopt_YM.
CC The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC Gag-polypeptides which may then be used to vaccinate against HIV
XX infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ Sequence 2306 BP; 529 A; 752 C; 730 G; 295 T; 0 other;

Query Match 82.2%; Score 2019.2; DB 21; Length 2306;
Best local Similarity 93.2%; Pred. No. 4.8e-243;
Matches 2150; Conservative 0; Mismatches 138; Indels 18; Gaps 3;

QY 170 GCGCAAGGAGGGCCACAGATGAGGACTGCACCGAGCGCCAGCGCAACTTCTTCGCG 229
Db 1 CGGCGCGGAGGACACCAATGAAGATTGCATGAGACAGACTAATTTCTTCGCG 60
QY 230 AGGACCTGGCCTTCCCCAGGGCAAGGCCGCGAGTTCCCGCAGCAGCAACCGCGCA 289
Db 61 AGGACCTGGCCTTCTGCAGGCAAGGCCGCGAGTTTCAGCAGCAGCAGCACCGCGCA 120
QY 290 ACAGCCCCACAGCCGGAGCTGCAGGTGCGGGG-----ACAAACCCCGCAGCGAG 343
Db 121 ACAGCCCCACCGCGGAGCTGCAGGTGTGGGGGCGGAGAACCAACAGCTGAGCGAG 180
QY 344 CCGGCGCGAGCGCCAGGCGACCCCTG-----AACTTCCCCAGATCACCTGTGGCAGC 397
Db 181 CCGGCGCGAGCGCCAGGCGACCGTGAGCTTCACTTCCCCAGATCACCTGTGGCAGC 240
QY 398 GCCCCTGGTGAGCATCAAGTGGCGGCCAGATCAAGAGGCCCTGTGTGACACCGCG 457
Db 241 GCCCCTGGTGACCATCAGGATCGGGCGCCAGCTCAAGAGGCGCTGTGTGACACCGCG 300
QY 458 CCGAGCACACCGTGTGTGAGCAGATGAGCCTGCCGCGCAAGTGAAGCCCAAGATGATCG 517
Db 301 CCGAGCACACCGTGTGTGAGCAGATGAGCCTGCCGCGCAAGTGAAGCCCAAGATGATCG 360
QY 518 CGCGCATCGCGGCTTTCATCAAGTGCGCCAGTAGCAGCCAGATCTGTGATCGATCTCG 577
Db 361 CGCGGATCGGGGCTTTCATCAAGTGCGCCAGTAGCAGCCAGATCTCGGATCTCGC 420
QY 578 GCAAGAGGCCATCGGCACCGTGTGATCGGCCACCCACCCCGTGAACATCATCGCGCGCA 637
Db 421 GCCAAGAGGCCATCGGCACCGTGTGTGGGCCCCACCCCGTGAACATCATCGCGCGCA 480
QY 638 ACATGCTGACCCAGCTGGGCTGCACCCCTGAACCTTCCCCATCAGCCCCATCGAGACCGTGC 697
Db 481 ACCTGCTGACCCAGATCGCTGCACCTGAACCTTCCCCATCAGCCCCATCGAGACCGTGC 540
QY 698 CCGTGAAGCTGAAGCCCGGATGAGCGGCCCGCAAGGTGAAGAGTGGGCCCTTGACCGAG 757
Db 541 CCGTGAAGCTGAAGCCCGGATGAGCGGCCCGCAAGGTGAAGAGTGGGCCCTTGACCGAG 600
QY 758 AGAAGATCAAGGCCCTGACCGCCATCTCTCGAGGAGATGGAAGGAGGCGCAAGATCACCA 817
Db 601 AGAAGATCAAGGCCCTGTGTGAGATCTGCACCGAGATGGAGAGGCGCAAGATCACCA 660
QY 818 AGATCGGCCCGAGAACCCCTTACAAACCCCGCTGTTCGCCATCAAGAGAGGACAGCA 877
Db 661 AGATCGGCCCGAGAACCCCTTACAAACCCCGCTGTTCGCCATCAAGAGAGGAGACAGCA 720

XX The present sequence is the coding sequence of a synthetic HIV
CC gag-reverse transcriptase expression cassette, FS(-).ProtMod.Ropt(+).
CC The Gag protein of HIV is needed for the assembly of virus-like
CC particles. In addition, the Gag protein is involved in many stages of the
CC HIV life cycle, including assembly, virion maturation after particle
CC release and early post-entry steps in viral replication. The expression
CC cassette may be used for the recombinant expression of HIV
CC gag-polypeptides which may then be used to vaccinate against HIV
CC infection and acquired immunodeficiency syndrome (AIDS).
XX
SQ Sequence 2312 BP; 533 A; 750 C; 732 G; 297 T; 0 other;

Query Match 81.3%; Score 1998.4; DB 21; Length 2312;
Best Local Similarity 92.9%; Pred. No. 1.8e-240;
Matches 2147; Conservative 0; Mismatches 141; Indels 24; Gaps 4;

QY 170 GCGCAAGGAGGCGCCACCATGAAAGACTGCACCGAGCGCGCCACCACTTCTCCGCG 229
DB 1 GCGGCGCGGAAGGACACCAATGAAGATTGCACCTGAGAGACAGGCTAATTTCTCCGCG 60
QY 230 AGGACCTGGCTTCCCGCAGGCAAGGCGCGAGTTCCCGAGCGAGCAACCGCGCA 289
DB 61 AGGACCTGGCTTCTGAGGCGAAGCGCGGAGTTGAGCGAGCGAGACCGCGCA 120
QY 290 ACAGCCCCACAGCGCGGAGCTGCAGTGCAGGCGGCGG-----ACAACCCCGCGAGCGAG 343
DB 121 ACAGCCCCACCGCGCGAGCTGCAGTGTGGGCGGCGGAGAAACAACAGCTGAGCGAGG 180
QY 344 CCGGCGCGAGCGCGAGGCGACCTG-----AACTTCCCCCAGATCACCTGTGGGAGC 397
DB 181 CCGGCGCGAGCGCGAGGCGACCGTGAAGTTCAACTTCCCCCAGATCACCTGTGGGAGC 240
QY 398 GCCCCTGGTGAGCATCAAGTGGCGCGCAGATCAAGAGGCGCTGCTGACACCGCGCG 457
DB 241 GCCCCTGGTGACCATCAGATCGCGCGCAGCTCAAGAGGCGGTGCTGACACCGCGCG 300
QY 458 CCGAGGACACCGTGTGAGGAGATGAGCTGCGCGGCAAGTGAAGCCCAAGATGATCG 517
DB 301 CCGAGGACACCGTGTGAGGAGATGAACCTGCGCGGCAAGTGAAGCCCAAGATGATCG 360
QY 518 CGGCATCGCGCGCTTCATCAAGTGGCGCAGTACGACCGAGTCTGATCAGATCTGCG 577
DB 361 CGGGATCGGGGCTTCATCAAGTGGCGCAGTACGACCGATCCCGGTGGAGATCTGCG 420
QY 578 GCAAGAAGGCCATCGGCACCGCTGTGATCGCGCCACCGCGTGAACATCATCGCGCGCA 637
DB 421 GCCAAGGCCATCGGCACCGCTGTGTTGGCGCCACCGCGTGAACATCATCGCGCGCA 480
QY 638 ACATGCTGACCGCTGGGCTGCACCTGAACTTCCCGCATCAGCCCGCATCGAGACCGTGC 697
DB 481 ACCTGCTGACCGCATCGGCTGCACTTCCCGTGAACCTTCCCGCATCAGCCCGCATCGAGCGGTGC 540
QY 698 CCGTGAAGCTGAAGCGCGCATGACGCGCCCAAGGTGAAGCTGAGCGAGTGAAGGAGGCGCAACATCACCA 817
DB 601 AGAAGATCAAGGCCCTGGTGGAGATCTGCACCGAGATGAGAGGAGGCGCAAGATCAGCA 660
QY 818 AGATCGCGCGCGAAGCCCTTACAAACCGCGGTGTTGCGCATCAAGAAGAGGAGCA 877
DB 661 AGATCGCGCGCGAAGCCCTTACAAACCGCGGTGTTGCGCATCAAGAAGAGGAGCA 720
QY 878 CCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACACCGCGGTGTTGCGCATCAAGAAGAGGAGCA 937
DB 721 CCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACACCGCGGTGTTGCGCATCAAGAAGAGGAGCA 780
QY 938 AGGTGCAAGCTGGGCGATCCCGACCGCGCGCTGGAAGAAGAGGAGGCTGACCGTGC 997
DB 781 AGGTGCAAGCTGGGCGATCCCGACCGCGCGCTGGAAGAAGAGGAGGCTGACCGTGC 840

QY 998 TGGACGTGGGCGACGCTACTTTCAGCGTCCCGCTGACGAGGAGCTTCCGCAAGTACACCG 1057
DB 841 TGGACGTGGGCGACGCTACTTTCAGCGTCCCGCTGACGAGGAGCTTCCGCAAGTACACCG 900
QY 1058 CTTTCAACATCCCGAGCTCAACAAGAGAGACCGCGCGCATCCGCTACCAAGTCAACAGTGC 1117
DB 901 CTTTCAACATCCCGAGCTCAACAAGAGAGACCGCGCGCATCCGCTACCAAGTCAACAGTGC 960
QY 1118 TGCCCCAGGGCTGGAAGGCGACCGCGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1177
DB 961 TGCCCCAGGGCTGGAAGGCGACCGCGCATCTTCCAGAGCAGCATGACCAAGATCCTGG 1020
QY 1178 AGCCCTTCCGCGCGCGCAACCGCGAGATCGTGTATCTACCA-----GGCCCCCTGTAGC 1231
DB 1021 AGCCCTTCCGCGCGCGCAACCGCGAGATCGTGTATCTACCAAGTACATGAGACCTGTAGC 1080
QY 1232 TGGGAGGAGACCTGAGATTCGCCAGCACCAGCGCGCAAGATCGAGGAGTGCACAAGCACC 1291
DB 1081 TGGGAGGAGACCTGAGATTCGCCAGCACCAGCGCGCAAGATCGAGGAGTGCGCCAGCACC 1140
QY 1292 TGCTCGGCTGGGCTTCACCAACCGCGCGCAAGAGCACCAGAGGAGCGCCCTTCTCTGT 1351
DB 1141 TGCTCGGCTGGGCTTCACCAACCGCGCGCAAGAGCACCAGAGGAGCGCCCTTCTCTGT 1200
QY 1352 CCAT-----CGAGCTGCACCGCGCAAGTGCAGCGCTGCAGCGCATCGAGCTGCCCGAGA 1405
DB 1201 GGATGGGCTACGAGCTGCACCGCGCAAGTGCAGCGCTGCAGCGCATCATGCTGCCCGAGA 1260
QY 1406 AGGAGAGTGGACCTGAAACGACATCCAGAAGCTGGTGGGCAAGTGAATCGGGCAGCG 1465
DB 1261 AGGAGAGTGGACCTGAAACGACATCCAGAAGCTGGTGGGCAAGTGAATCGGGCAGCG 1320
QY 1466 AGATCTACCGCGCATCAAGTGGCGCGCAGCTGTGCAAGCTGCTGCGCGCGCGCAAGCGCC 1525
DB 1321 AGATCTACCGCGCATCAAGTGAAGCTGTGCAAGCTGCTGCGCGCGCGCAAGCGCGCC 1380
QY 1526 TGACCGACATCGTGCCTCGCTGACCGAGGAGCGAGCTGGAGTGGCGAGAACCGCGAGA 1585
DB 1381 TGACCGAGTGTATCCCTCGCTGACCGAGGAGCGAGCTGGAGTGGCGCGAGAACCGCGAGA 1440
QY 1586 TCTTCGCGAGCGCTGACCGCGTGTACTAGACCGCAAGAGCTGTGGTGGCGCGAGA 1645
DB 1441 TCCTGAAGAGCGCGCTGACCGAGTGTACTAGACCGCAAGAGCTGTGGTGGCGCGAGA 1500
QY 1646 TCCAGAAGCAGGCGCACCGACCTGAGCTACACAGATCTACAGAGCGCTTTCAGAGAC 1705
DB 1501 TCCAGAAGCAGGCGCAGGCGAGTGGAGCTTACAGATCTTACAGAGCGCTTTCAGAGAC 1560
QY 1706 TGAAGACCGCAAGTACCGCAAGATTCGCACCGCGCCACACCAACCGTGAAGCAGCTGA 1765
DB 1561 TGAAGACCGCAAGTACCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1620
QY 1766 CCGAGGCGCTGCAGAGATTCGCATGAGAGAGATCGTGTATGAGGCGCAAGACCGCAAGT 1825
DB 1621 CCGAGGCGCTGCAGAGATTCGCATGAGAGAGATCGTGTATGAGGCGCAAGATTCGCAAGT 1680
QY 1826 TCCGCTTCCCATCCAGAAGGAGACCTGAGAGCTGTGACCGCTTACTTGGCAGGCGCA 1885
DB 1681 TCAAGCTGCCCATCCAGAAGGAGACCTGAGAGCGCTGTGGAGCGCTGTGGTGGAGGCGCA 1740
QY 1886 CCTGGATCCCGAGTGGAGTTCGTGAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1945
DB 1741 CCTGGATCCCGAGTGGAGTTCGTGAACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
QY 1946 TGGAGAGGAGCGCATCATCGCGCGCGCGAGACTTCTTACGTGGAGCGCGCGCGCGCGCGCG 2005
DB 1801 TGGAGAGGAGCGCATCATCGCGCGCGCGAGACTTCTTACGTGGAGCGCGCGCGCGCGCGCG 1860
QY 2006 AGACCAAGATCGGCAAGCGCGCTACGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2065
DB 1861 AGACCAAGCTGGGCAAGCGCGCGCTACGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1920
QY 2066 TGACCGAGACCAACCAACAGAGAGCCGAGCTGCGAGGCGCATCCAGCTGCGCGCGCGCGCG 2125

Db 1921 TCAGCGACCAACCAAGAGCGAGCTGCAGGCGCATCCACCTGGCCCTGCAGGACA 1980
QY 2126 GCGGACGAGGTGAACATCGTACCGACAGCAGTACGCCCTGGGCGCATCATCAGGCC 2185
Db 1981 GCGGCCCTGGAGGTGAACATCGTACCGACAGCAGTACGCCCTGGGCGCATCATCAGGCC 2040
QY 2186 AGCCGCAAGAGCGAGCGAGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAAGG 2245
Db 2041 AGCCGCAAGAGCGAGCGAGCTGGTGAACAGCATCATCGAGCAGCTGATCAAGAAGG 2100
QY 2246 AGAAGGTGTACCTGAGCTGGTGGCCGCCCAAGAGGSCATCGCGGCAACGACAGATCG 2305
Db 2101 AGAAGGTGTACCTGAGCTGGTGGCCGCCCAAGAGGSCATCGCGGCAACGACAGTGG 2160
QY 2306 ACAGCTGGTGAAGGAGGATCGGCAAGGTGCTGTTCCTGGAGCGGATCGATGGCGGCA 2365
Db 2161 ACAAGCTGGTGAAGGAGGATCGGCAAGGTGCTGTTCCTGGAGCGGATCGATGGCGGCA 2220
QY 2366 TCGTGATCTACCACTACATGAGCAGCCTGTACGTGGGAGCGGCGGCTAGGATCGATT 2425
Db 2221 TCGTGATCTACCACTACATGAGCAGCCTGTACGTGGGAGCGGCGGCTAGGATCGATT 2280
QY 2426 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2457
Db 2281 AAAAGCTTCCCGGGCTAGCACCGGTGAATTC 2312

RESULT 8

ABK91622
ID ABK91622 standard; DNA; 9788 BP.

XX
AC ABK91622;
XX

DT 14-AUG-2002 (first entry)
XX

DE Modified HIV protein-encoding plasmid DNA #174.
XX

XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX

OS Synthetic.
XX

XX WO200232943-A2.
XX

XX 25-APR-2002.
XX

XX 14-AUG-2001; 2001WO-US25721.
XX

XX 14-AUG-2000; 2000US-225097P.
XX

PR 14-NOV-2000; 2000US-252115P.
PR

PR 28-MAR-2001; 2001US-279257P.
PR

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX

PA (CHAD/) CHADABARTI B K.
PA

XX Nabel GJ, Huang Y;
XX

XX NPI; 2002-452382/48.
XX

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection
XX

XX Disclosure; Page 785-788; 794pp; English.
PS
XX

CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX

SQ Sequence 9788 BP; 2377 A; 2817 C; 2695 G; 1899 T; 0 other;

Query Match 81.1%; Score 1992.4; DB 24; Length 9788;
Best Local Similarity 91.2%; Pred. No. 8.4e-240;

Matches 2181; Conservative 0; Mismatches 181; Indels 30; Gaps 5;

QY 14 TGGCCGAGGCGCATGAGCGAGCGCCACCAACATCTGTATCGAGCGCAACATTCA 73
Db 2561 TGGCCGAGGCGCATGAGCGAGGTGAACAACCAACATCATGTGACGCGCAACTGCA 3020

QY 74 AGGCGCCCAAGCGCATCATCAAGTGTCTTAACCTCGCGAAGGAGGCGCACATCGCCGCA 133
Db 3021 AGGCGCCCAAGCGCACCATCAAGTGTCTTAACCTCGCGAAGGAGGCGCACATCGCCGCA 3080

QY 134 ACTGCCGCGCCCGCCGCAAGAGGCTGTGGAAGTGCAGGAGGAGGCGCCACAGATGA 193
Db 3081 ACTGCCGCGCCCGCCGCAAGAGGCTGTGGAAGTGCAGGAGGAGGCGCCACAGATGA 3140

QY 194 AGGACTGCACCGAGCGCGAGGCCCAACTTCTTCCGCGAGGAGCTGTGCCCTTCCCGCAGGGCA 253
Db 3141 AGGACTGCACCGAGCGCGAGGCCCTAA-----TAGGGAAGATCTGGCCTTCCCAACAAGGA 3194

QY 254 AGGCGCGCGAGTTCCCGAGGAGGAGAGAACCGCGCAACAGCCCGCCACAGCGCGGAGTGC 313
Db 3195 AGGCGAGGGAATTTCTTTCAGAGCAGACACAGAGCAACAGCCCGCCACAGAGAGAGTTC 3254

QY 314 AGTGCGCGG-----CGACAACCCCGCGAGGCGCGCGGCGAGCGCGCAGGCA--- 364
Db 3255 AGGTTTGGGGAAGAGACAACACTCCCTCTCAGAGGAGGAGCGCGATAGACAAGAACTG 3314

QY 365 ---CCCTGAACCTCCCGCAGATCACCTCTGGCGAGCGCCCTTGGTGAAGTGAAGTGG 421
Db 3315 TATCCTTTAGTCTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCAACAATAAGATAG 3374

QY 422 GCGGCGAGATCAAGAGGCGCTGTGGACACCGCGCGGAGACACCGTGTGGAGGAGA 481
Db 3375 GGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGGAGACACCGTGTGGAGGAGA 3434

QY 482 TGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGGCGCATCGCGGCTTCATCAAGG 541
Db 3435 TGAACCTGCCCGCGCTGGAGCCCAAGATGATCGGCGCATCGCGGCTTCATCAAGG 3494

QY 542 TGCGCCAGTACGACCATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCGCCGTGC 601
Db 3495 TGGGCCAGTACGACCATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCGCCGTGC 3554

QY 602 TGATCGGCGCCACCCCGTGAACATCATCGCGCGCAACATCTGACCCAGCTGGGCTGCA 661
Db 3555 TGGTGGCGCCACCCCGTGAACATCATCGCGCGCAACATCTGACCCAGATCGGCTGCA 3614

QY 662 CCCTGAACCTCCCGCATCGAGCGCGTGGCGCATCGAGCGCGGCTTCATCAAGG 721
Db 3615 CCCTGAACCTCCCGCATCGAGCGCGTGGCGCATCGAGCGCGGCTTCATCAAGG 3674

QY 722 ACGCCCGCAAGTGAAGCAGTGGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCCGCA 781
Db 3675 ACGCCCGCAAGTGAAGCAGTGGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCCGCA 3734

QY 782 TCTGGAGGAGATGGAGAGGAGGCGCAAGATCAACAAGATCGGCCCGCGAGACCCCTACA 841
Db 3735 TCTGGAGGAGATGGAGAGGAGGCGCAAGATCAACAAGATCGGCCCGCGAGACCCCTACA 3794

QY 842 ACACCCCGCTGTCCGCCATCAAGAGAAGAGGAGACAGCACCAAGTGGCGCAAGCTGGTGGACT 901
Db 3795 ACACCCCGCTGTCCGCCATCAAGAGAAGAGGAGACAGCACCAAGTGGCGCAAGCTGGTGGACT 3854

QY 902 TCCGCGAGCTGAACAAGCGCCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCCCACC 961
Db TCCGCGAGCTGAACAAGCGCCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCCCACC 3914
QY 962 CCGCGGCTGAAGNAGNAGAGCGTGCCTGCTGGAGCTGGCGACGCCCTACTTCA 1021
Db TCCGCGGCTGAAGNAGNAGAGCGTGCCTGCTGGAGCTGGCGACGCCCTACTTCA 3974
QY 1022 GCGTGCCTTGGAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCAAGCATCAACA 1081
Db GCGTGCCTTGGAGGAGCTTCCGCAAGTACACCGCTTCCACCATCCCAAGCATCAACA 4034
QY 1082 ACAGAGCCCCGGGATCCGCTACCAAGTACAAAGTGTCTGCCCGAGGCTTGAAGGCGAGCC 1141
Db ACAGAGCCCCGGGATCCGCTACCAAGTACAAAGTGTCTGCCCGAGGCTTGAAGGCGAGCC 4094
QY 1142 CCAGCATCTCCAGAGAGATGACCAAGATCCCTGGAGCCCTTCCGCGCCCGCAACCCCG 1201
Db CCAGCATCTCCAGAGAGATGACCAAGATCCCTGGAGCCCTTCCGCGCCCGCAACCCCG 4154
QY 1202 AGATCGTGATACCA-----GGCCCCCTGTACGTGGGCGAGCATCGGATCGGCC 1255
Db AGATCGTGATACCA-----GGCCCCCTGTACGTGGGCGAGCATCGGATCGGCC 4214
QY 1256 AGCAGCGGCGAAGATCGAGAGCTGCGCAAGCACTGTCTGCGTGGGGCTTACCAACC 1315
Db AGCAGCGGCGAAGATCGAGAGCTGCGCAAGCACTGTCTGCGTGGGGCTTACCAACC 4274
QY 1316 CCGACAAGAGCACCAGAGAGCCCTTCTTCTGCCCAT-----CGAGTGCACCCCG 1369
Db CCGACAAGAGCACCAGAGAGCCCTTCTTCTGCCCAT-----CGAGTGCACCCCG 4334
QY 1370 ACAAGTGAACCGTGCAGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCTGGAACGACA 1429
Db ACAAGTGAACCGTGCAGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCTGGAACGACA 4394
QY 1430 TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 1489
Db TCCAGAAGCTGGTGGGCAAGCTGAACCTGGGCGACCCAGATCTACCCCGCATCAAGTGC 4454
QY 1490 GCCAGCTGTCAGCTGCTGGGGGCGCCAGGCGCTGACCGACATCGTGGCCCTGACCG 1549
Db GCCAGCTGTCAGCTGCTGGGGGCGCCAGGCGCTGACCGACATCGTGGCCCTGACCG 4514
QY 1550 AGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCTGACCGCG 1609
Db AGGAGCGGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCTGACCGCG 4574
QY 1610 TGACTAGACCCAGCAAGAGCTGTGTGGCGGAGATCCAGAGCAGGGCCACGACAGT 1669
Db TGACTAGACCCAGCAAGAGCTGTGTGGCGGAGATCCAGAGCAGGGCCACGACAGT 4634
QY 1670 GGACCTACAGATCTACAGAGCCCTTCAAGACCTGAGACCGGCAAGTACGCCAAGA 1729
Db GGACCTACAGATCTACAGAGCCCTTCAAGACCTGAGACCGGCAAGTACGCCAAGA 4694
QY 1730 TGGCGACCCGCCACCAACAGCTGAAGCAGCTGACCGAGGCGGTGACAGAGTGCCTCA 1789
Db TGGCGACCCGCCACCAACAGCTGAAGCAGCTGACCGAGGCGGTGACAGAGTGCCTCA 4754
QY 1790 TGGAGAGCATCTGTGTGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAAGAGA 1849
Db TGGAGAGCATCTGTGTGGGCAAGACCCCAAGTTCCGCCCTGCCCATCCAGAAGAGA 4814
QY 1850 CTTGGGAGACTGTGTGACCGACTACTGGCAGGCCACCTGGATCCCGGAGTGGAGTTCG 1909
Db CTTGGGAGACTGTGTGACCGACTACTGGCAGGCCACCTGGATCCCGGAGTGGAGTTCG 4874
QY 1910 TGAACACCCCCCTTCTGGTGAAGCTGTGTGTTACAGCTGGAGAGGAGCCCATCATCGGCG 1969
Db TGAACACCCCCCTTCTGGTGAAGCTGTGTGTTACAGCTGGAGAGGAGCCCATCATCGGCG 4934
QY 1970 CCGAGACCTTCTACGTGGAGCGCGCCGCCAACCCGCGAGACCAAGATCGGCAAGCGCGCT 2029

Db 4935 CCGAGACCTTCTACGTGGAGCGCGCCCAACCCGAGACCAAGCTGGCAAGCGCGCT 4994
QY 2030 ACGTGACCGACCGGGCGCGGAGAGATCGTGAAGCTTGACCGAGACCAACCAAGAGA 2089
Db ACGTGACCGACCGGGCGCGGAGAGATCGTGAAGCTTGACCGAGACCAACCAAGAGA 5054
QY 2090 CCGAGCTCGAGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCTGTA 2149
Db CCGAGCTCGAGCCATCCAGCTGGCCCTGCAGGACAGCGGCGAGGTGAACATCTGTA 5114
QY 2150 CCGAGCCAGTACGCTGGGCGATCATCCAGGCGCCAGCCGACAGAGCGAGCGAGC 2209
Db CCGAGCCAGTACGCTGGGCGATCATCCAGGCGCCAGCCGACAGAGCGAGCGAGC 5174
QY 2210 TGGTGAACCATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGGTGC 2269
Db TGGTGAACCATCATCGAGCAGCTGATCAAGAGGAGAGTGTACCTGAGCTGGGTGC 5234
QY 2270 CCGCCCAAGGGCATCGGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCATCC 2329
Db CCGCCCAAGGGCATCGGCGCAACGAGCAGATCGACAAGCTGGTGAGCGCGCGCATCC 5294
QY 2330 GCAAGTGTCTTCTCTGGACGCGATCGATGGCGCATCGTGTATCTACAGTA 2381
Db GCAAGTGTCTTCTCTGGACGCGATCGACAAGCGCCAGGAGGAGGACGAGAA 5346

RESULT 9
ABK91617
ID ABK91617 standard; DNA; 9169 BP.
XX
AC ABK91617;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #169.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
OS Synthetic.
XX
PN WO200232943-A2.
XX
PD 25-APR-2002.
XX
XX 14-AUG-2001; 2001WO-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-225115P.
PR 28-MAR-2001; 2001US-279257P.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
(CHAD/) CHADABARTI B K.
PA
PI Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
XX
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure; Page 769-772; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9169 BP; 2227 A; 2668 C; 2505 G; 1769 T; 0 other;

Query Match 80.5%; Score 1978.4; DB 24; Length 9169;
Best Local Similarity 91.1%; Pred. No. 4.6e-238;
Matches 2178; Conservative 0; Mismatches 186; Indels 28; Gaps 6;

QY 14 TGGCCGAGGCATGAGCGAGCCACAGCGCCCAACATCTCTGATGCGAGCGAGCAACTTCA 73
DB 2967 TGGCCGAGGCATGAGCGAGCGATGCGACACACCAATCATGATGCGAGCGGCACTTCA 3026
QY 74 AGGCCCCCAAGCGATCATCAAGTGTTCACCTTCACTGCGGCAAGGAGGCCACATCGCCGCA 133
DB 3027 AGGCCAGAGCGCAT---CAAAGTGTTCACCTGCGGCAAGGAGGCCACCTGCGCCGCA 3083
QY 134 ACTGCCGCGCCCGCCCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCCACCATGA 193
DB 3084 ACTGCCGCGCCCGCTGCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCCACCATGA 3143
QY 194 AGGACTGCACCGAGCGCGAGCCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCGCAGGCA 253
DB 3144 AGGACTGCACCGAGCGACAGCTAA-TTTTTTAGGGAAGATCTGGCCCTTCCCAAGGGA 3202
QY 254 AGGCCCGGAGTCCCGAGCGAGGAGAGAACCGCGCCACACCGCCACCGCGGAGCTGC 313
DB 3203 AGGCCAGGAAATTTCTTCAGAGCAGACACAGAGCCACACCGCCACCGAGAGGCTTC 3262
QY 314 AGGTGCGCGG-----CGACAAACCCCGCAGGAGCGCGCGAGCGCGCAGGCA--- 364
DB 3263 AGGTTGGGAAGAGACAACTCCCTCTCAGAGGAGGAGCGCGATAGACAGGAAGTCTG 3322
QY 365 ---CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGTGG 421
DB 3323 TATCCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACATAAAGATAG 3382
QY 422 CGGCCAGATCAAGGAGCGCTCTGTGACACCGCGCGCGAGCACCGCTGCTGAGGAGA 481
DB 3383 GGGCCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGAGCACCGCTGCGAGGAGA 3442
QY 482 TGAGCTGCGCGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCAAG 541
DB 3443 TGAACCTGCGCGCGCTGGAAGCCCAAGATGATCGCGGCGATCGCGGCTTCATCAAGG 3502
QY 542 TGGCCAGTACGACAGATCCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCGTGC 601
DB 3503 TGGCCAGTACGACAGATCCTGATCGAGATCTGCGGCCACAAAGGCCATCGGCACCGTGC 3562
QY 602 TGATCGGCCCGCCCGCTGACATCATCGGCCGACATGCTGACCCAGCTGGGCTGCA 661
DB 3563 TGTGGGCGCCCGCCCGCTGAACATCATCGGCGCGCAACCTGCTGACCCAGATCGGCTGCA 3622
QY 662 CCCTGAACTTCCCGCATCAGCCCGCATGAGACCGCTGCGCGTGAAGCTGAAGCCCGGCGATGG 721
DB 3623 CCCTGAACTTCCCGCATCAGCCCGCATGAGACCGCTGCGCGTGAAGCTGAAGCCCGGCGATGG 3682
QY 722 ACGGCCCAAGTGAAGAGTGGCCCTGTGCGGAGGAGAGATCAAGGCCCTGACCCGCCA 781
DB 3683 ACGGCCCAAGTGAAGAGTGGCCCTGTGCGGAGGAGAGATCAAGGCCCTGAGTGGAGA 3742
QY 782 TCTCGAGGAGATGGAAGAGGCGGAGATCACCAAGATCGGCCCGCGAGAACCCCTTACA 841
DB 3743 TCTGACCGAGATGGAAGAGGCGGAGATCAGCAAGATCGGCCCGCGAGAACCCCTTACA 3802
QY 842 ACACCCCGCTTCCCGCATCAAGAGAGGAGACAGCACAAGTGGCGCAAGTGGTGACT 901
DB 3803 ACACCCCGCTTCCCGCATCAAGAGAGGAGACAGCACAAGTGGCGCAAGTGGTGACT 3862

QY 902 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCCCACC 961
DB 3863 TCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGACGCTGGGATCCCCCACC 3922
QY 962 CCGCCGCGCTGAAGAAGAAAGAGCGTACCGTGTGGAGCTGGGGGAGCGCTACTTCA 1021
DB 3923 CCGCCGCGCTGAAGAAGAAAGAGCGTACCGTGTGGAGCTGGGGGAGCGCTACTTCA 3982
QY 1022 GCGTCCCTGGAGCGAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGCATCAACA 1081
DB 3983 GCGTCCCTGGACAGGACTTCCGCAAGTACACCGCTTACCATCCCGAGCATCAACA 4042
QY 1082 ACAGACCCCGCGCATCCGCTACAGTACAACTGCTGCCCCAGGGCTGGAAGGGCAGCC 1141
DB 4043 ACAGACCCCGCGCATCCGCTACAGTACAACTGCTGCCCCAGGGCTGGAAGGGCAGCC 4102
QY 1142 CCAGCATCTTCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGAGCGCCGCAACCCG 1201
DB 4103 CCAGCATCTTCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGAGCGCCGCAACCCG 4162
QY 1202 AGATCGTGTATACCA-----GGCCCCCTGTACGTGGGCGAGCACTGGAGATCGGCC 1255
DB 4163 ACATCGTGTATACCACTGACCATGACCACTGTACGTGGGCGAGCACTGGAGATCGGCC 4222
QY 1256 AGACCCCGCGAGATCGAGAGCTGCGCAAGCACTGCTGCGCTGGGGCTTCAACACCC 1315
DB 4223 AGACCCCGCGAGATCGAGAGCTGCGCAAGCACTGCTGCGCTGGGGCTTCAACACCC 4282
QY 1316 CGCACAAGAGCACCAGAAGGAGCCCTTCTCTGCCAT-----CGAGCTGCACCCCG 1369
DB 4283 CGCACAAGAGCACCAGAAGGAGCCCTTCTCTGCCAT-----CGAGCTGCACCCCG 4342
QY 1370 ACAAGTGGACCGTGCAGCCCATCGAGCTGCGGAGAGAGAGCTGGAACCGTGAACGACA 1429
DB 4343 ACAAGTGGACCGTGCAGCCCATCGTGTGCGCGAGAGAGAGCTGGAACCGTGAACGACA 4402
QY 1430 TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCGCAGCCAGATCTACCCCGGCGATCAAGTGC 1489
DB 4403 TCCAGAAGCTGGTGGGCAAGCTGAACTGGGCGCAGCCAGATCTACCCCGGCGATCAAGTGC 4462
QY 1490 GCCAGCTGTCAAGCTGTGCGCGCGCGCCAGGCGCTGACCGACATGCTGCGCCCTGACCG 1549
DB 4463 GCCAGCTGTCAAGCTGTGCGCGCGCGCCAGGCGCTGACCGAGCTGTGCGCCCTGACCG 4522
QY 1550 AGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCGCCGTGCAGCGCG 1609
DB 4523 AGGAGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGAGGAGCGCCGTGCAGCGCG 4582
QY 1610 TGTAAGTACGAGCTGAGAGAGCTGGTGGCGGAGATCCAGAGCGAGCGCGCGAGT 1669
DB 4583 TGTAAGTACGAGCTGAGAGAGCTGGTGGCGGAGATCCAGAGCGAGCGCGCGAGT 4642
QY 1670 GGACCTACGAGCTTACAGGAGCGCTTCAAGAACCTGGAAGCGCGCAAGTACGCGCAAGA 1729
DB 4643 GGACCTACGAGCTTACAGGAGCGCTTCAAGAACCTGGAAGCGCGCAAGTACGCGCGCA 4702
QY 1730 TGCACACCGCCCGCACCAACAGCTGAAAGAGCTGACCGAGCGCGCTGCAAGAGATCGGCA 1789
DB 4703 TGAAGGCGCGCCCGCACCAACAGCTGAAAGAGCTGACCGAGCGCGCTGCAAGAGATCGGCA 4762
QY 1790 TGGAGAGCATCTGATCTGGGCGAAGACCCCAAGTTCGCGCTGCCATCCAGAGGAGA 1849
DB 4763 CCGAGAGCATCTGATCTGGGCGAAGACCCCAAGTTCGCGCTGCCATCCAGAGGAGA 4822
QY 1850 CCTGGGAGCTGCTGGAGCGGCTTCTGCGAGCGCACTGATCCCGAGTGGGAGTTCG 1909
DB 4823 CCTGGGAGCTGCTGGAGCGGCTTCTGCGAGCGCACTGATCCCGAGTGGGAGTTCG 4882
QY 1910 TGAACACCGCCCGCTGGTGAAGTGTGTACAGCTGGAGAGAGGCCCATCATCTCGCGG 1969
DB 4883 TGAACACCGCCCGCTGGTGAAGTGTGTACAGCTGGAGAGAGGCCCATCATCTCGCGG 4942
QY 1970 CCGAGACCTTCTACGTGGACCGCGCCCGCAACCGCGAGACCAAGATCGCAAGCGCGCT 2029

|||||
Db 4943 CCGAGACCTTCTAGTGGACGGCGCCCAACCGGAGACCAAGCTGGGCAAGSCCGCT 5002
QY 2030 ACGTGACGACCGGGCGCGCAGAGATCGTGGAGCTGACCGAGACCAACACAGAGA 2089
Db 5003 ACGTGACGACCGGGCGCGCAGAGATCGTGGAGCTGACCGAGACCAACACAGAGA 5062
QY 2090 CCGAGCTGCAGGCGATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCCTGA 2149
Db 5063 CCGAGCTGCAGGCGATCCAGCTGGCCCTGCAGGACAGCGGAGGTGAACATCCTGA 5122
QY 2150 CCGAGACCAAGTACGCGCTGGGATCATCCAGGCCCGCCAGCCGACAGAGGAGAGCGAGC 2209
Db 5123 CCGAGACCAAGTACGCGCTGGGATCATCCAGGCCCGCCAGCCGACAGAGGAGAGCGAGC 5182
QY 2210 TGGTGAACACAGATCATCGACGCTGTATCAAGAAGGAGAGGTGTACTGAGCTGGGTGC 2269
Db 5183 TGGTGAACACAGATCATCGACGCTGTATCAAGAAGGAGAGGTGTACTGAGCTGGGTGC 5242
QY 2270 CCGCCCAAGAGGCGATCGGCGGCAACGACGAGATCGACAGCTGGTGAGCAAGGGCATCC 2329
Db 5243 CCGCCCAAGAGGCGATCGGCGGCAACGACGAGGTGGAGCGGCTGGTGAGCGCCGGCATCC 5302
QY 2330 GCAAGGTGCTGTTCTCTGGAGCGGATCGATGGCGGCATCGTGATCTACAGTA 2381
Db 5303 GCAAGGTGCTGTTCTCTGGAGCGGATCGACAAAGGCCAGGAGGACACAGAA 5354

RESULT 10
ABK91612
ID ABK91612 standard; DNA: 9194 BP.
AC ABK91612;
XX 14-AUG-2002 (first entry)
XX Modified HIV protein-encoding plasmid DNA #164.
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
XX Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
XX acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX Synthetic.
XX W0200232943-A2.
XX 25-APR-2002.
XX 14-AUG-2001; 2001W0-US25721.
XX 14-AUG-2000; 2000US-225097P.
XX 14-NOV-2000; 2000US-252115P.
XX 28-MAR-2001; 2001US-279257P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (CHAD/) CHADABARTI B K.
XX Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
XX its encoded protein, useful as vaccines for genetic or protein
XX immunisation for acquired immunodeficiency syndrome or human
XX immunodeficiency virus infection -
XX Disclosure; Page 753-756; 794pp; English.
XX The invention relates to a nucleic acid molecule encoding a modified HIV
XX (human immunodeficiency virus) protein. The nucleic acid molecule and its
XX encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
XX genetic or protein immunisation to a host, respectively. In particular
XX these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9194 BP; 2232 A; 2672 C; 2523 G; 1767 T; 0 other;

Query Match 80.4%; Score 1975; DB 24; Length 9194;
Best Local Similarity 90.9%; Pred. No. 1.2e-237;
Matches 2177; Conservative 0; Mismatches 190; Indels 28; Gaps 6;

QY 14 TGGCCGAGGCGCATAGCCAGG-CCACCAGCGCCACATCCTGTATGAGCGCAGCAACT 70
Db 2967 TGGCCGAGGCGCATAGCCAGGTCACCAACAGCGCCACCATCATGTATGAGCGCGCAACT 3026
QY 71 TCAAGGGCGCCCAACGCGATCATCAAGTCTTCAACTGCGCAAGGAGGCGCCACATCGGCC 130
Db 3027 TCCGCAACCAAGCGCAAGATCGTGAAGTCTTCAACTGCGCAAGGAGGCGCCACCGCCC 3086
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGGCTCTGGAAGTGGGCAAGGAGGCGCCACCAAGA 190
Db 3087 GCAACTGCGCGCGCCCGCCGCAAGAGGCTCTGGAAGTGGGCAAGGAGGCGCCACCAAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCGCCAGCCCAACTTCTTCCGCGAGGACCTGCCTTCCCCCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGCGAGCGCTAA-TTTTATGAGGAAGATCTGGGCTTCCCACAG 3205
QY 251 GCAAGGCGCGGAGTTCCCGACGAGCAGCAACCGCGCCCAACAGCGCCCAACAGCGCGCAGC 310
Db 3206 GGAAGGCGAGGAAATTTCTTACAGCAGCAGCAGCCACAGCGCCCAACAGCAGAGAGAGC 3265
QY 311 TGCAGGTGGCGG-----CGACAACCGCCGACGAGCGCGCGCGCGCGCGCGCGCGCA 364
Db 3266 TTCAGGTTGGGGAAGAGACAACAACCTCTCAGAAGCAGGAGCGCATAGACAAGGAA 3325
QY 365 -----CCCTGAACCTTCCCGCAGATCACCTGTGGCAGCGCGCCCTGTGTAGCATCAGG 418
Db 3326 CTGTATCTTTAGTCTTCCCTCAGATCACTTTTGGCAGCGACCCCTCGTCAACAATAAGA 3385
QY 419 TGGCGGCGCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 3386 TAGGGGCGCAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 3445
QY 479 AGATGAGCGCTGCCCGGCAAGTGGAAAGCCCAAGATGATGCGGCGGATCGGCGGCTTCATCA 538
Db 3446 AGATGAACCTGCCCGGCGCTGGAAGCCCAAGATGATGCGGCGGATCGGCGGCTTCATCA 3505
QY 539 AGGTGCGCGCAGTACGACCATCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
Db 3506 AGGTGCGGCGAGTACGACCATCTGTATCGAGATCTGCGGCGCAAGAGGCCATCGGCACCG 3565
QY 599 TGTGATGCGCGCCCGCCCGCTGAACATCATCGCGCGCGCAACATGTGACCCAGCTGGGCT 658
Db 3566 TGTGTTGGCGCCCGCCCGCTGAACATCATCGCGCGCGCAACATGTGACCCAGATCGGCT 3625
QY 659 GCACCTGAACTTCCCGCATCGCCCGCTGAGACCGCTGCCCGTGAAGCTGAAGCGCCGCA 718
Db 3626 GCACCTGAACTTCCCGCATCGCCCGCTGAGACCGCTGCCCGTGAAGCTGAAGCGCCGCA 3685
QY 719 TGGACGCGCCCGCGTGAAGCAGTGGCGCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3686 TGGACGCGCCCGCGTGAAGCAGTGGCGCTGACCGAGGAGAGATCAAGGCCCTGACCG 3745
QY 779 CCATCTGCGAGGAGATGGAGAAGGAGGCAAGATCACCAAGATCGGCCCGCGAGAACCCCT 838
Db 3746 AGATCTGACCGAGATGGAGAAGGAGGCAAGATCAAGAACATCGGCCCGCGAGAACCCCT 3805
QY 839 ACAACACCGCGTGTGGCCATCAAGAGAGAGGAGCAGCACCAGTGGCGCGAAGTGGTGG 898
Db 3806 ACAACACCGCGTGTGGCCATCAAGAGAGAGGAGCAGCACCAGTGGCGCGAAGTGGTGG 3865

QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGAGACTTCTGGGAGGTGCAGCTGGGATCCCC 958
Db 3866 ACTTCCGCGAGCTGAACAAGCGCACCCAGAGACTTCTGGGAGGTGCAGCTGGGATCCCC 3925
QY 959 ACCCGCGCGCTGAAGAAGAAAGAGCTGACCGTGTCTGGAGCTGGGGGAGCGCTTACT 1018
Db 3926 ACCCGCGCGCTGAAGAAGAAAGAGCTGACCGTGTCTGGAGCTGGGGGAGCGCTTACT 3985
QY 1019 TCAGCTGCCCCCTGGAGGAGGACTTCGCAAGTACACCGCTTTCACCATCCCCAGATCA 1078
Db 3986 TCAGCTGCCCCCTGGAGGAGGACTTCGCAAGTACACCGCTTTCACCATCCCCAGATCA 4045
QY 1079 ACAACGAGACCCCGGATCCGCTACAGTACAACTGTCTGCCCGAGGGGTGGAAGGCA 1138
Db 4046 ACAACGAGACCCCGGATCCGCTACAGTACAACTGTCTGCCCGAGGGGTGGAAGGCA 4105
QY 1139 GCCCCAGCATCTCCAGAGCAGATGACCAAGATCTCTGGAGCCCTTCCCGCGCCGAACC 1198
Db 4106 GCCCCAGCATCTCCAGTGCAGCATGACCAAGATCTCTGGAGCCCTTCCCGCAAGCAACC 4165
QY 1199 CCAGATCTGTATACCA-----GGCCCCCTGTACGTGGGAGGAGCTGGAGATCG 1252
Db 4166 CCAGATCTGTATACCAAGTACATGACCACTGTACGTGGGAGGAGCTGGAGATCG 4225
QY 1253 GCCAGACCCCGGAGATCGAGGAGCTGCGCAAGCAGTGTCTGCGTGGGGCTTCACCA 1312
Db 4226 GCCAGACCCCGGAGATCGAGGAGCTGCGCAAGCAGTGTCTGCGTGGGGCTTCACCA 4285
QY 1313 CCCCCAGAAAGACGACCAAGAGGCCCCCTTCTGCCCCAT-----CGAGCTGCACC 1366
Db 4286 CCCCCAGAAAGACGACCAAGAGGCCCCCTTCTGCCCCATGGGCTACGAGCTGCACC 4345
QY 1367 CCAGCAAGTGGACCGTGCAGCCCATCGAGCTGCCCGAGAGGAGAGCTGGACCGTGAACG 1426
Db 4346 CCAGCAAGTGGACCGTGCAGCCCATCGTGTCTGCCCGAGAGGAGAGCTGGACCGTGAACG 4405
QY 1427 ACATCAGAAAGCTGTGGGCAAGCTGAATCTGGGCGCAGCCAGATCTACCCCGGATCAAGG 1486
Db 4406 ACATCAGAAAGCTGTGGGCAAGCTGAATCTGGGCGCAGCCAGATCTACCCCGGATCAAGG 4465
QY 1487 TGGCGCAGCTGTCAAGCTGTCTGGCGGCGCCAAAGCCCTGACCGGACATCTGTCCTCTGA 1546
Db 4466 TGGCGCAGCTGTCAAGCTGTCTGGCGGCGCCAAAGCCCTGACCGGAGGTGGTCCCTCTGA 4525
QY 1547 CCAGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGCGGAGCCCGTGCACG 1606
Db 4526 CCAGAGGCGCGAGCTGGAGCTGGCGGAGAACCGCGAGATCTCTGAGAGGCGCGTGCACG 4585
QY 1607 GCCTGTACTAGACCCCGCAGCAAGGACTGTGTGGCGGAGATCTCAGAGAGCGGCGACGAC 1666
Db 4586 GCCTGTACTAGACCCCGCAGCAAGGACTGTGTGGCGGAGATCTCAGAGAGCGGCGACGAC 4645
QY 1667 AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 4646 AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 4705
QY 1727 AGATGGGACCGCCGACCAACAGCGTGAAGCAGCTGACCGAGGCGGTGCAGAAAGATCG 1786
Db 4706 GCATGAAGGCGCCGACCAACAGCGTGAAGCAGCTGACCGAGGCGGTGCAGAAAGATCG 4765
QY 1787 CCATGGAGAGCATCTGATCTGGGGCAAGACCCCAAGTGTCCGCTTCCGATCCCAAGG 1846
Db 4766 CCACCGAGAGCATCTGATCTGGGGCAAGACCCCAAGTGTCCGCTTCCGATCCCAAGG 4825
QY 1847 AGACCTGGGAGACTGTGTGACCGACTACTGGCAGGCCACCTGGATCCCCGAGTGGAGT 1906
Db 4826 AGACCTGGGAGGCTGTGTGACCGAGTACTGGCAGGCCACCTGGATCCCCGAGTGGAGT 4885
QY 1907 TCGTGAACACCCCCCTCGGTGAAGCTGTGTACCAAGCTGGGAGAGGAGGCCCATCATCG 1966
Db 4886 TCGTGAACACCCCCCTCGGTGAAGCTGTGTACCAAGCTGGGAGAGGAGGCCCATCATCG 4945
QY 1967 GCGCCGAGACTTCTTAGTGGAGCGCGCGCCAAACCGCGAGACCAAGATCGGCAAGGCGG 2026

Db 4946 GCGCCGAGACTTCTAGCTGGAGCGCGCCCAACCGGAGACCAAGCTGGCAAGGCCG 5005
QY 2027 GCTACGCTGACCGCGCGCGCGGAGAGATCGTGAAGCTGACCGAGACCAACACAGA 2086
Db 5006 GCTACGCTGACCGCGCGCGCGGAGAGAGTGTGTGCCCCCTGACCGAGACCAACACAGA 5065
QY 2087 AGACCCGAGCTGCGAGGCGCATCCAGCTGGCCCTGCGAGGACGCGGAGGTGAACATCG 2146
Db 5066 AGACCGAGCTGCGAGGCGCATCCACCTGGCCCTGCGAGGACGCGGCTGGAGTGAACATCG 5125
QY 2147 TGACCGAGACCCAGTAGTACGCCCTGGGCATCATCCAGGCCAGCCCGCAGAGGAGAGCG 2206
Db 5126 TGACCGAGACCCAGTAGTACGCCCTGGGCATCATCCAGGCCAGCCCGCAGAGGAGAGCG 5185
QY 2207 AGCTGTGTGAACAGAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2266
Db 5186 AGCTGTGTGAACAGAGATCATCGACGAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 5245
QY 2267 TGCCCGCCCAAGAGGCGCATCGCGGCGCAACGAGCAGATCGACAGCTGGTGAAGCGCA 2326
Db 5246 TGCCCGCCCAAGAGGCGCATCGCGGCGCAACGAGCAGGTGGAGCGCTGGTGAAGCGCGCA 5305
QY 2327 TCAGCAAGGTGTCTTCTCGACGCGCATCGATGGCGGATCGTGGGATCTTACAGTA 2381
Db 5306 TCAGCAAGGTGTCTTCTCGACGCGCATCGACAGGCCAGGCCAGGAGGAGAGAGAA 5360

RESULT 11

ABK91611
ID ABK91611 standard; DNA; 9194 BP.

XX ABK91611;

DT 14-AUG-2002 (first entry)

XX Modified HIV protein-encoding plasmid DNA #163.

XX HIV; human immunodeficiency virus; gene: ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX Synthetic.

XX WO200232943-A2.

XX 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

XX 14-NOV-2000; 2000US-252115P.

XX 28-MAR-2001; 2001US-279257P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (CHAD/) CHADABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure; Page 750-753; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ

Sequence 9194 BP; 2232 A; 2671 C; 2524 G; 1767 T; 0 other;

Query Match 80.3%; Score 1973.4; DB 24; Length 9194;
Best Local Similarity 90.9%; Pred. No. 1.9e-237;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;
QY 14 TGGCCGAGGCCATGAGCCAGG---CCACAGCGCCAAATCTCTGATGCGAGCCAGCAACT 70
DB 2967 TGGCCGAGGCCATGAGCCAGGTCGACCAAGAGCCGACCATGATGCGAGCGGGCAACT 3026
QY 71 TCAAGGSCCCCAAGCGCATCATCAAGTGCTTCAACTGCGGCAAGAGGGCCACATCGCCC 130
DB 3027 TCCGCAACCAAGCGCATGCTGAAGTGCTTCAACTGCGGCAAGAGGGGCCACACCGCCC 3086
QY 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGGCCACCA 190
DB 3087 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGGCCACCA 3146
QY 191 TGAAGGACTGCAACGAGCGCAGGCCAACTTCTTCGCGAGGAGCTGGCCCTCCCGCCAGG 250
DB 3147 TGAAGGACTGCAACGAGCGCAGGCCAACTTCTTCGCGAGGAGCTGGCCCTCCCGCCAGG 3205
QY 251 GCAAGGSCCGCGAGTTCCCGCAGGAGCAGCAACCGCGCCACAGCCCGCCGCGAGC 310
DB 3206 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACAGAGCCACAGCCCGCCACCAAGAGAGC 3265
QY 311 TCGAGGTGCGGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCA 364
DB 3266 TTCAGGTTTGGGGAAGAGACAACACTCCCTCTCAGAAGCAGGAGCGCGATAGACAAGGA 3325
QY 365 -----CCCTGAACTTCCCGCAGATACCTCTGTGCGAGCGCCCTCTGGTGAGATCAAGG 418
DB 3326 CTGTATCTTTAGTTCCTTCATCATCTCTTGGCAGCAGCCCTCTGCTGATCAATTAAGA 3385
QY 419 TGGCGGCCAGATCAAGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGCGCG 478
DB 3386 TAGGGGCCAGCTGAAGGAGCGCTTCTAGACACCGCGCGCGCGCGCGCGCGCGCGCG 3445
QY 479 AGATGACCTTGGCGGGAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 3446 AGATGAACCTTGGCGGCGCTGGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 3505
QY 539 AGGTGCGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 598
DB 3506 AGGTGCGCCAGTACGACAGATCTGATCGAGATCTGCGGCAAGAGGCCATCGGCACCG 3565
QY 599 TGTGATCGGCGCCACCGCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 3566 TGTGTTGGCGCCCGCCCGCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCT 3625
QY 659 GCACCTTGAATTCCTCCATCAGCCCATCGAGACCGTGGCCGTGAAGTGAAGCCCGGCA 718
DB 3626 GCACCTTGAATTCCTCCATCAGCCCATCGAGACCGTGGCCGTGAAGTGAAGCCCGGCA 3685
QY 719 TGGAGGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 3686 TGGAGGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGGTGG 3745
QY 779 CCATCTCGAGGAGATGGAAGAGGAGGCAAGATCAACAGATCGCGCGCGCGCGCGCGCG 838
DB 3746 AGATCTGACCGGAGTGAAGAGGAGGCAAGATCAACAGATCGCGCGCGCGCGCGCGCG 3805
QY 839 ACAACACCCCGGTTTGGCCATCAAGAAAGAGGACAGCACCAGTGGCGCAAGCTGGTGG 898
DB 3806 ACAACACCCCGGTTTGGCCATCAAGAAAGAGGACAGCACCAGTGGCGCAAGCTGGTGG 3865

QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCC 958
DB 3866 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGAGCTGGGCATCCCCC 3925
QY 959 ACCGCGCGGCTGTAAGAAGAAGAAGAGCGTGAACGCTGTGGACGTGGCGACGCTACT 1018
DB 3926 ACCGCGCGGCTGTAAGAAGAAGAAGAGCGTGAACGCTGTGGACGTGGCGACGCTACT 3985
QY 1019 TCAGGCTGCCCTGACGAGGAGACTTCCGCAAGTACACGCCCTTACCATCCCGACATCA 1078
DB 3986 TCAGGCTGCCCTGACGAGGAGACTTCCGCAAGTACACGCCCTTACCATCCCGACATCA 4045
QY 1079 ACAAGAGAGACCCCGCGCATCCCTACCACTACAAGTGTGCCCGAGGCTGGAAGGCA 1138
DB 4046 ACAAGAGAGACCCCGCGCATCCCTACCACTACAAGTGTGCCCGAGGCTGGAAGGCA 4105
QY 1139 GCCCAGCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTTCCGCGCGCCCAACC 1198
DB 4106 GCCCAGCATCTTCCAGTGCAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAAGACC 4165
QY 1199 CCGAGATCGTGATCTACCA-----GGCCCGCTGTAGTGGGCGAGCACCTGGAGATCG 1252
DB 4166 CCGAGATCGTGATCTACCACTACATGGAGACCTGTACGTGGGCGAGCACCTGGAGATCG 4225
QY 1253 GCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTACCA 1312
DB 4226 GCAGCACCGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGCTGGGCTTACCA 4285
QY 1313 CCCCAGCAAGAGACCCAGAGGAGCGCCCTTCTCGGCCAT-----CGAGCTGCAACC 1366
DB 4286 CCCCAGCAAGAGACCCAGAGGAGCGCCCTTCTCGGCCAT-----CGAGCTGCAACC 4345
QY 1367 CCGACAAGTGGACGTGCGAGCCCATCGAGCTGCGCGAGAGAGAGCTGGACCGTGAACG 1426
DB 4346 CCGACAAGTGGACGTGCGAGCCCATCGTGTGCGCGAGAGAGAGCTGGACCGTGAACG 4405
QY 1427 ACATCCAGAGCTGTGGCGAAGCTGAACTGGGCGCAGCAGATCTACCCCGCATCAAGG 1486
DB 4406 ACATCCAGAGCTGTGGCGAAGCTGAACTGGGCGCAGCAGATCTACCCCGCATCAAGG 4465
QY 1487 TCGCGCAGCTGTGCAAGCTGTGCGCGCGCAAGCGCTTACGACATCTGCGCCCTGA 1546
DB 4466 TCGCGCAGCTGTGCAAGCTGTGCGCGCGCAAGCGCTTACGAGGTGCGCCCTGA 4525
QY 1547 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCCCGTGCAGG 1606
DB 4526 CCGAGAGCGCGAGCTGGAGCTGGCGGAGAACCGGAGATCTGCGCGAGCCCGTGCAGG 4585
QY 1607 GCGTGTACTTACACCCCGCAGAGGAGCTGTGGCGGAGATCCAGAGCAGGCGCACGACC 1666
DB 4586 GCGTGTACTTACACCCCGCAGAGGAGCTGTGCGCGAGATCCAGAGCAGGCGCACGACC 4645
QY 1667 AGTGAACCTTACAGATCTTACAGGAGCGCTTCAAGAAGCTGAAGACCGCGCAAGTACGCCA 1726
DB 4646 AGTGAACCTTACAGATCTTACAGGAGCGCTTCAAGAAGCTGAAGACCGCGCAAGTACGCCA 4705
QY 1727 AGATGGCACCCCGCCACACCAAGAGCTGAAGCAGCTGACCGAGCGCGCTGCGAAGATCG 1786
DB 4706 GCATGAAGGCGCGCCACACCAAGAGCTGAAGCAGCTGACCGAGCGCGCTGCGAAGATCG 4765
QY 1787 CCATGGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 1846
DB 4766 CCACCGAGAGCATCTGATCTGGGCAAGACCCCAAGTTCGCGCTGCCCATCCAGAGG 4825
QY 1847 AGACCTGGGAGACCTTGGTGAGCCGACTACTGGCAGCGCCACCTGGATCCCGAGTGGGAGT 1906
DB 4826 AGACCTGGGAGCGCTTGGTGAGCCGAGTACTGGCAGCGCCACCTGGATCCCGAGTGGGAGT 4885
QY 1907 TCGTGAACACCCCGCCCTTGGTGAAGCTGTGTACAGCTGAGAGAGGAGCGCCATCATCG 1966
DB 4886 TCGTGAACACCCCGCCCTTGGTGAAGCTGTGTACAGCTGAGAGAGGAGCGCCATCATCG 4945
QY 1967 GCGCGGAGACCTTCTACGTGGAGCGCGCGCCCAACCGGAGACCAAGATCGCAAGGCGCG 2026

Db 4946 GCGCCGAGACCTTCTAGCTGAGCGCGCCGCCAACCGCGAGACCAAGCTGGCAAGGCCG 5005
QY 2027 GCTACGTTGACCGCGCGCGCGCAGAGATCGTGAGCCTGACCGAGACCAACACAGA 2086
Db 5006 GCTACGTTGACCGCGCGCGCGCAGAGATCGTGAGCCTGACCGAGACCAACACAGA 5065
QY 2087 AGACCGAGCTGACCGCGCGCGCGCAGAGATCGTGAGCCTGACCGAGACCAACACAGA 2146
Db 5066 AGACCGAGCTGACCGCGCGCGCGCAGAGATCGTGAGCCTGACCGAGACCAACACAGA 5125
QY 2147 TGACCGAGACCGAGTACCGCGCGCGCGCAGAGATCGTGAGCCTGACCGAGACCAACACAGA 2206
Db 5126 TGACCGAGACCGAGTACCGCGCGCGCGCAGAGATCGTGAGCCTGACCGAGACCAACACAGA 5185
QY 2207 AGCTGGTGAACCGAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACTGAGCTGGG 2266
Db 5186 AGCTGGTGAACCGAGATCATCGAGCAGCTGATCAAGAGGAGAGAGTGTACTGAGCTGGG 5245
QY 2267 TGCCCGCGCCACAAAGGGGATCGCGCGCGCGCGCAGAGATCGAGCAAGGGCA 2326
Db 5246 TGCCCGCGCCACAAAGGGGATCGCGCGCGCGCGCAGAGATCGAGCAAGGGCA 5305
QY 2327 TCCGCAAGGTGCTTCTGCGAGCGGATCGCGCGCGCGCGCAGAGATCGAGCAAGGGCA 2381
Db 5306 TCCGCAAGGTGCTTCTGCGAGCGGATCGCGCGCGCGCGCAGAGATCGAGCAAGGGCA 5360

RESULT 12

ABK91613
ID ABK91613 standard; DNA; 12411 BP.
XX
AC ABK91613;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #165.
XX
KW HIV; human immunodeficiency virus; gene; ds: circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN W0200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001W0-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADABARTI B K.
XX
XX Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
PS Disclosure; Page 756-760; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX

SQ Sequence 12411 BP; 2938 A; 3733 C; 3476 G; 2364 T; 0 other;

Query Match 80.3%; Score 1973.4; DB 24; Length 12411;
Best Local Similarity 90.9%; Pred. No. 1.9e-237;
Matches 2176; Conservative 0; Mismatches 191; Indels 28; Gaps 6;

QY 14 TGCCGAGCGCCATGAGCAGG---CCACCAGCGCCACATCTCTGATCGACGCGCAACT 70
Db 2967 TGCCGAGCGCCATGAGCAGGTCACCAACAGCGCCACCATCATGATCGACGCGGCACT 3026
QY 71 TCAGGCGCCCAAGCGCATCATCAAGTGTTCAACTCGCGCAAGGAGGGCCACATCGCCC 130
Db 3027 TCCGCAACCGACGCAAGATCGTGAAGTGTTCAACTCGCGCAAGGAGGGCCACACGCC 3086
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGGGCAAGAGGGCCACACAGA 190
Db 3087 GCAACTGCCCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGGGCAAGAGGGCCACACAGA 3146
QY 191 TGAAGGACTGCACCGAGCGCCGCAACTTCTTCCCGAGGAGCTTGGCTTCCCGCCAGG 250
Db 3147 TGAAGGACTGCACCGAGCGAGAGGCTTAA-TTTTATGAGAGATCTGGCTTCCACACAG 3205
QY 251 GCAAGCGCCCGCGAGTTTCCCGAGCGAGCAACCGCGCCAAACAGCGCCCGCCAGCGAGC 310
Db 3206 GGAAGCGCAGGGAATTTCTTTCAGAGCAGACAGAGCAACAGCGCCCGCCAGAGAGAGC 3265
QY 311 TGCAGGTGCGCGG-----CGACAACCGCGAGCGAGCGCGCGCGCGCGCGAGGGCA 364
Db 3266 TGCAGGTGCGCGGAGAGAGCAACCACTTCTCAGAGCAGGAGCGCGCGCGCGCGAGGAA 3325
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
Db 3326 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTGCTCACAATAAGA 3385
QY 419 TGCGCGCGCAGATCAAGGAGGCGCTTGTGACACCGCGCGCGCGAGCACCGCTGCGAGG 478
Db 3386 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGCGAGCACCGCTGCGAGG 3445
QY 479 AGATGAGCCTGCCCGCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGGGCGCTTCATCA 538
Db 3446 AGATGAGCCTGCCCGCGCGCGCTGGAAGCCCAAGATGATCGCGCGCATCGGGCGCTTCATCA 3505
QY 539 AGGTGCGCCAGTACGACCGAGATCTGTATCGAGATCTCGCGCAAGAGCGCATCGGCAACG 598
Db 3506 AGGTGCGCCAGTAGGACCGAGATCTGTATCGAGATCTCGCGCCACAGCGCATCGGCAACG 3565
QY 599 TGTGATCGCGCGCGCGCGCTGAACATCATCGCGCGCGCGCAACATGCTGACCGAGCTGGGCT 658
Db 3566 TGTGATCGCGCGCGCGCGCTGAACATCATCGCGCGCGCGCAACATGCTGACCGAGCTGGCT 3625
QY 659 GCACCTGAACTTCCCGCATCGACCGCATCGAGACCGTGCCTGCGAGCTGAAGCGCGGCA 718
Db 3626 GCACCTGAACTTCCCGCATCGACCGCATCGAGACCGTGCCTGCGAGCTGAAGCGCGGCA 3685
QY 719 TGCACGCGCCCAAGGTGAAGCAGTGGCGCTTACCGAGGAGAGATCAAGCGCGCTGACCG 778
Db 3686 TGCACGCGCCCAAGGTGAAGCAGTGGCGCTTACCGAGGAGAGATCAAGCGCGCTGCTGG 3745
QY 779 CCATCTGCGAGGAGATGGAGAGGAGGCAAGATCAACAGATCGCGCGCGCGAGAACCCCT 838
Db 3746 AGATCTGCACCGAGATGGAGAGGAGGCAAGATCAAGAGATCGCGCGCGCGAGAACCCCT 3805
QY 839 ACAACACCGCGCTTTCGCCCATCAAGAGAGAGGACAGCACCAAGTGGCGGCAAGCTGGTG 898
Db 3806 ACAACACCGCGCTTTCGCCCATCAAGAGAGAGGACAGCACCAAGTGGCGGCAAGCTGGTG 3865

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

SQ Sequence 9785 BP; 2360 A; 2843 C; 2731 G; 1851 T; 0 other;

Query Match	80.2%;	Score 1970;	DB 24;	Length 9785;
Best Local Similarity	90.9%;	Pred. No. 5.1e-237;		
Matches 2174; Conservative	0;	Mismatches 185;	Indels 33;	Gaps 6

[illegible]

Db 4938 CCGAGACCTTCTAGCTGGACGGCCGCCCAACCGCGAGACCAAGCTGGGCAAGCCGGCT 4997
QY 2030 ACCTGACCGACCGCGCGGCGGAGAGATCGTGAGCCCTGACCGAGACCAACCAAGCAAGA 2089
Db 4998 ACCTGACCGACCGCGGCGGAGAGATCGTGAGCCCTGACCGAGACCAACCAAGCAAGA 5057
QY 2090 CCGAGCTGACAGGCGATCCAGCTGGCCCTGCGAGGACGCGGCGAGGAGTGAACATCGTGA 2149
Db 5058 CCGAGCTGACAGGCGATCCACCTGGCCCTGCGAGGACGCGGCGTGGAGTGAACATCGTGA 5117
QY 2150 CCGACAGCCAGTACGCTGGGATCATCCAGGCGCCAGCGCGGAGAGAGAGAGAGAGAG 2209
Db 5118 CCGACAGCCAGTACGCTGGGATCATCCAGGCGCCAGCGCGGAGAGAGAGAGAGAGAG 5177
QY 2210 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGAGTGTACTGAGCTGGGTGC 2269
Db 5178 TGGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAGAGTGTACTGAGCTGGGTGC 5237
QY 2270 CCGCCACAGAGGCGATCGGCGGCAACGAGCAGATCGACAAAGCTGGTGAGCAAGGCGATCC 2329
Db 5238 CCGCCACAGAGGCGATCGGCGGCAACGAGCAGTGGAGCGGCGCTGGTGAGCGCGGCGATCC 5297
QY 2330 GCAGGCTGCTTCTCGAGCGGCGATCGATGGCGCATCGTGATCTACAGTA 2381
Db 5298 GCAAGGTGCTTCTCGAGCGGCGATCGACAAAGCGCGGAGGAGGAGGAGGAGGAGGAG 5349

RESULT 14
ABK91619
ID ABK91619 standard; DNA; 9167 BP.
XX AC ABK91619;
DT 14-AUG-2002 (first entry)
DE Modified HIV protein-encoding plasmid DNA #171.
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX Synthetic.
OS
XX WO200232943-A2.
PN 25-APR-2002.
XX
XX 14-AUG-2001; 2001WO-US25721.
XX
XX 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USHS) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADRABARTI B K.
XX
XX Nabel GJ, Huang Y;
XX WPI; 2002-452382/48.
XX
XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
PT immunodeficiency virus infection -
XX
XX Disclosure; Page 775-778; 794pp; English.
XX
XX The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.
XX
SQ Sequence 9167 BP; 2225 A; 2668 C; 2517 G; 1757 T; 0 other;
Query Match 80.08; Score 1966.6; DB 24; Length 9167;
Best Local Similarity 90.7%; Pred. No. 1.4e-236;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCCCAACTCTGATGAGCGCAGCAACT 70
Db 2964 TGGCCGAGGCGCATGAGCCAGGCGCCACACAGCGCCACCATGATGAGCGCGCAACT 3023
QY 71 TCAGAGGCGCCCAAGCGCATCATCAAGTCTTCAACTGCGGAAGAGGCGCCACATPCGCC 130
Db 3024 TCCGCAACACGCGCAAGATCGTGAAGTCTTCAACTGCGGAAGAGGCGCCACACCGCC 3083
QY 131 GCAACTGCGCGCGCCCGCCGCAAGAGAGGCTGCTGGAAGTGGCGGCAAGAGGCGCCAC 190
Db 3084 GCAACTGCGCGCGCCCGCCGCAAGAGAGGCTGCTGGAAGTGGCGGCAAGAGGCGCCAC 3143
QY 191 TGAAGGAGTGCACCGAGCGCCGCGCAACTTCTTCCGCGAGGACCTGGCTTCCCGCAGG 250
Db 3144 TGAAGGAGTGCACCGAGCGCGCAAGCTAA-----TAGGGAAGATCTGGCTTCCGCAAG 3197
QY 251 GCAAGGCGCGGAGTTCGCCAGCGAGCAGAACCGCCCAACAGCGCCACCGCGCGAGC 310
Db 3198 GGAAGGCGCGGAGTTCGCCAGCGAGCAGAACCGCCCAACAGCGCCACCGCGCGAGC 3257
QY 311 TGCAGGTGCGGG-----CGACAACCGCCCGAGGAGCGCGCCCGAGCGCCAGGCGCA 364
Db 3258 TTCAGGTTGGGGAAGAGACAACACTCCCTCTCAGAAAGAGAGCGCATAGACAAGGAA 3317
QY 365 -----CCCTGAATCTCCCGCAGATCACCTGTGCGAGCGCCCTGGTGAGCATCAAGG 418
Db 3318 CTGTATCTTTAGCTTCCCTCAGATCACTTTGGCAGCGACCCCTGCTCAATAAAGA 3377
QY 419 TGGCGCGCGAGATCAAGAGGCGCCCTGCTGGACACCGCGCGGAGACACCTGTGTGAGG 478
Db 3378 TAGGGGCGCAGCTGAAGGAGGCGCTTCTAGACACCGCGCGGAGACACCTGTGTGAGG 3437
QY 479 AGATGAGCTTCCCGGCAAGTGGAGCCCAAGATGATCGGGGCGATCGGGCGCTTCATCA 538
Db 3438 AGATGAACCTGCGCGCGCTGGAAGCGCCCAAGATGATCGGGGCGATCGGGCGCTTCATCA 3497
QY 539 AGGTGCGCGAGTACGACCGAGATCTGATCGAGATCTGCGCAAGAGGCCATCGGCACCG 598
Db 3498 AGGTGCGCGAGTACGACCGAGATCTGATCGAGATCTGCGGCGCAAGGCCATCGGCACCG 3557
QY 599 TGCTGATCGGCGCCCGCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGTGGGCT 658
Db 3558 TGCTGCTGGGCGCCCGCCCGTGAACATCATCGCGCGCAACCTGCTGACCCAGATCGGCT 3617
QY 659 GCACCTGAACCTTCCCGATCAGCCCGATCGAGACCGTGGCGGCGATGAAGTGAAGCGCGCA 718
Db 3618 GCACCTGAACCTTCCCGATCAGCCCGATCGAGACCGTGGCGGCGATGAAGTGAAGCGCGCA 3677
QY 719 TGGAGCGCGCCCAAGTGAAGCAGTGGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 3678 TGGAGCGCGCCCAAGTGAAGCAGTGGCGCCCTGACCGAGGAGAGATCAAGGCCCTGTTGG 3737
QY 779 CCATCTCGAGGAGATGGAGAGGAGGCGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 838
Db 3738 AGATCTGCACCGAGATGGAGAGGAGGCGCAAGATCAACCAAGATCGGCCCGCGAGAACCCCT 3797
QY 839 ACAACACCGCGTTCGCCATCAAGAGAGGAGACAGCAACCAAGTGGCGGCAAGCTGGTGG 898
Db 3798 ACAACACCGCGTTCGCCATCAAGAGAGGAGACAGCAACCAAGTGGCGGCAAGCTGGTGG 3857

QY 899 ACTTCGCGAGCTGAACAAGCGCACCCAGAGACTTCTGGAGGTGACGCTGGGCATCCCC 958
Db 3858 ACTTCGCGAGCTGAACAAGCGCACCCAGAGACTTCTGGAGGTGACGCTGGGCATCCCC 3917
QY 959 ACCCGCGCGCTGAAGAGAAGAAAGAGCTGACCGTGTGGAGCTGGGGGACGCTACT 1018
Db 3918 ACCCGCGCGCTGAAGAGAAGAAAGAGCTGACCGTGTGGAGCTGGGGGACGCTACT 3977
QY 1019 TCAGCGTGCCTCGGAGGAGCTTCGCAAGTACACCGCTTCACATCCCGCAGATCA 1078
Db 3978 TCAGCGTGCCTCGGAGGAGCTTCGCAAGTACACCGCTTCACATCCCGCAGATCA 4037
QY 1079 ACAACGAGACCCCGCATCCGTTACCAAGTACAAAGTGTGCCCCAGGGCTGGAAGGCA 1138
Db 4038 ACAACGAGACCCCGCATCCGTTACCAAGTACAAAGTGTGCCCCAGGGCTGGAAGGCA 4097
QY 1139 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCTTGGAGCCCTTCCGCGCCCGCAACC 1198
Db 4098 GCGCCAGCATCTCCAGAGCAGATGACCAAGATCTTGGAGCCCTTCCGCGCCCGCAACC 4157
QY 1199 CCGAGATCGTGTATCA-----GGCCCGCTGTACGTGGGCGAGCGACTTGGAGATCG 1252
Db 4158 CCGATCGTGTATCACTACAGTATGACCACTGTACGTGGGCGAGCGACTTGGAGATCG 4217
QY 1253 GCAGACCCGCGCAAGATCGAGAGCTGCGCAAGCAGCACTGTGCGCTGGGGTTCAACA 1312
Db 4218 GCAGACCCGCGCAAGATCGAGAGCTGCGCAAGCAGCACTGTGCGCTGGGGTTCAACA 4277
QY 1313 CCGCCGACAGACGACCAAGAGGCCCCCTTCTGCCCCAT-----CGAGCTGCACC 1366
Db 4278 CCGCCGACAGAGACGACCAAGAGGCCCCCTTCTGCCCCAT-----CGAGCTGCACC 4337
QY 1367 CCGACAAAGTGGACCGTGCACGCCATCGAGCTGCCGAGAAAGGAGAGCTGGACCGTGAACG 1426
Db 4338 CCGACAAAGTGGACCGTGCACGCCATCGTGTGCCGAGAAAGGAGAGCTGGACCGTGAACG 4397
QY 1427 ACATCCAAAGTGTGGGGCAAGCTGAACCTGGGGCCAGCAGATCTACCCCGGATCAAGG 1486
Db 4398 ACATCCAAAGTGTGGGGCAAGCTGAACCTGGGGCCAGCAGATCTACCCCGGATCAAGG 4457
QY 1487 TGGCGCAGCTGTCAGAGCTGCTCGCGGCGCCAAAGCCCTGACCGAGATCGTCCCTGA 1546
Db 4458 TGGCGCAGCTGTCAGAGCTGCTCGCGGCGCCAAAGCCCTGACCGAGGTGGTGGCCCTGA 4517
QY 1547 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTCGCGAGGCGCGTGACG 1606
Db 4518 CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCGCGAGATCTCTCGCGAGGCGCGTGACG 4577
QY 1607 GCGTGTACTACGACCCAGCAAGACCTGTGTGGCCGAGATCCAGAGCAGGGCCAGACC 1666
Db 4578 GCGTGTACTACGACCCAGCAAGACCTGTGTGGCCGAGATCCAGAGCAGGGCCAGGGCC 4637
QY 1667 AGTGGACCTACAGATCTACAGAGCCCTTCAAGACCTGAAGACCGCGCAAGTACGCCA 1726
Db 4638 AGTGGACCTACAGATCTACAGAGCCCTTCAAGACCTGAAGACCGCGCAAGTACGCCC 4697
QY 1727 AGATGGCACCCGCCACCAACGACGTGAAGCAGCTGACCGAGCGCGTGCAGAAATCG 1786
Db 4698 GCATGAGGGCGGCCACCAACGACGTGAAGCAGCTGACCGAGCGCGTGCAGAAATCG 4757
QY 1787 CCATGGAGACATCGTGTATCTGGGGCAAGACCCCAAGTTCCGCTGCCATCCAGAGG 1846
Db 4758 CCACCGAGACATCGTGTATCTGGGGCAAGACCCCAAGTTCCAGCTGCCCATCCAGAGG 4817
QY 1847 AGACCTGGGAGACCTGCTGACCGACTACTTGGCAGGCCACCTGGATCCCGAGTGGAGT 1906
Db 4818 AGACCTGGGAGGCGCTGGTGACCGAGTACTTGGCAGGCCACCTGGATCCCGAGTGGAGT 4877
QY 1907 TCGTGAACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGCCCATCATCG 1966
Db 4878 TCGTGAACACCCCGCTGGTGAAGCTGTGTACCAAGCTGGAGAGGAGCCCATCATCG 4937
QY 1967 GCGCCGAGACCTTCTAGTGGAGCGGCGCGCCCAACCGCGAGACCAAGATCGGCAAGGCG 2026

Db 4938 GCGCCGAGACCTTCTAGTGGAGCGGCGCCCAACCGCGAGACCAAGCTGGCAAGGCCG 4997
QY 2027 GCTACGTGACCGACCGGGCGCGCAAGATCGTGAAGCTGAGCGAGACCAACCAAGC 2086
Db 4998 GCTACGTGACCGACCGGGCGCGCAAGATCGTGAAGCTGAGCGAGACCAACCAAGC 5057
QY 2087 AGACCGAGCTGACAGGCGCATCCAGCTGGCCCTGCAGGACAGCGAGGTGAACATCG 2146
Db 5058 AGACCGAGCTGACAGGCGCATCCACCTGSCCCTGCAGGACAGCGCCTGGAGTGAACATCG 5117
QY 2147 TGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCCCCAGCCGACAAAGAGGAGCG 2206
Db 5118 TGACCGACAGCCAGTACGCGCTGGGCATCATCCAGGCCCCAGCGCCGACAAAGAGGAGCG 5177
QY 2207 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACTGTAGCTGGG 2266
Db 5178 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAGGAGAGGTGTACTGTAGCTGGG 5237
QY 2267 TGCCCGCCCAAGAGGCGCATCGCGGCAACGAGCAGATCGCAAGCTGCTGAGCAAGGCA 2326
Db 5238 TGCCCGCCCAAGAGGCGCATCGCGGCAACGAGCAGGTGGAGCGGCTGCTGAGCGCGCA 5297
QY 2327 TCGCAAGGTGCTGCTTCTGACGCGCATCGATGGCGCATCGTGTATACAGTA 2381
Db 5298 TCGCAAGGTGCTGCTTCTGACGCGCATCGCAAGCGCCAGGAGGAGGACACAGAA 5352

RESULT 15
ABK91614
ID ABK91614 standard; DNA; 9170 BP.
XX
AC ABK91614;
XX
DT 14-AUG-2002 (first entry)
XX
DE Modified HIV protein-encoding plasmid DNA #166.
XX
KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;
KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;
KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.
XX
OS Synthetic.
XX
PN WO200232943-A2.
XX
PD 25-APR-2002.
XX
PF 14-AUG-2001; 2001WO-US25721.
XX
PR 14-AUG-2000; 2000US-225097P.
PR 14-NOV-2000; 2000US-252115P.
PR 28-MAR-2001; 2001US-279257P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (CHAD/) CHADABARTI B K.
XX
PI Nabel GJ, Huang Y;
XX
XX WPI; 2002-452382/48.
DR
XX
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and
PT its encoded protein, useful as vaccines for genetic or protein
PT immunisation for acquired immunodeficiency syndrome or human
XX immunodeficiency virus infection -
PS Disclosure; Page 760-763; 794pp; English.
XX
CC The invention relates to a nucleic acid molecule encoding a modified HIV
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its
CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for
CC genetic or protein immunisation to a host, respectively. In particular
CC these are useful for ameliorating the symptoms of acquired

CC immunodeficiency syndrome (AIDS) or HIV infection and generating an
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent
CC compositions are useful for treating or preventing HIV infections or
CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding
CC plasmid DNA sequences of the invention.

XX
SQ Sequence 9170 BP; 2225 A; 2669 C; 2519 G; 1757 T; 0 other;

Query Match 80.0%; Score 1966.6; DB 24; Length 9170;
Best Local Similarity 90.7%; Pred. No. 1.4e-236;
Matches 2173; Conservative 0; Mismatches 189; Indels 33; Gaps 6;

QY 14 TGGCCGAGGCGATGAGCCAGG---CCACCAGCGCAACATCTTCATCGAGCGAGCAACT 70

DB 2967 TGGCCGAGGCGATGAGCCAGGTCACCAACAGCGCCACCATCATGATGAGCGCGCAACT 3026

QY 71 TCRAAGGCGCCAGCGCATCATCAAGTGTCTTAACCTCGGCAAGAGGGCCACATCGCC 130

DB 3027 TCCGCAACCAACGCAAGATCGTGAAGTGTCTTAACCTCGGCAAGAGGGCCACATCGCC 3086

QY 131 GCAACTCGCGCGCCCGCCCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 190

DB 3087 GCAACTCGCGCGCCCGCCCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACACAGA 3146

QY 191 TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACTTGGCCCTTCCCCAGG 250

DB 3147 TGAAGGACTGCACCGAGCGCCAGGCCAACTTCTTCCGCGAGGACTTGGCCCTTCCCCAGG 3200

QY 251 GCRAAGGCGCGGAGTTCCCGAGCGAGCAGAGAACCGCGCCACACAGCCCAACGCGCGAGC 310

DB 3201 GGAAGGCGCGGAGTTTCTTACAGAGCAGACACAGAGCAACAGCCCAACAGAGAGAGC 3260

QY 311 TGCAAGTGCAGG---CGACAACCCCGCAGCGAGGCGCGCGCGCGCGCAGGCA 364

DB 3261 TTCAAGTTTGGGGAAGAGAACAACTTCCCTCTCAGAAGCAGGAGCGGATAGACAAGAA 3320

QY 365 -----CCCTGAATCTCCCGCAGATCACTTGTGGAGCGCCCTGTGTGAGCATCAAG 418

DB 3321 CTGTATCTTTCCTTCAGATCACTTCTTGGCAGCGACCCCTCTGCACATAAGA 3380

QY 419 TGGCGCGCGAGATCAAGAGGCGCTGTGTGACACCGCGCGCGAGCAGCATGTGTGAGG 478

DB 3381 TAGGGGCGAGCTGAAGAGGCGCTTCTAGACACCGCGCGCGAGCAGCATGTGTGAGG 3440

QY 479 AGATGAGCTTCCCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCTTCATCA 538

DB 3441 AGATGAGCTTCCCGCGCGCTTGGAAAGCCCAAGATGATCGCGCGCATCGCGCTTCATCA 3500

QY 539 AGGTGCGCCAGTAGACACAGATCTGTATCGAGATCTTGGGCAAGAGGCCATCGGCACCG 598

DB 3501 AGGTGCGCCAGTAGACACAGATCTGTATCGAGATCTTGGGCAAGAGGCCATCGGCACCG 3560

QY 599 TGTGATCGGCGCCACCGCGTGAACATCATCGCGCGCACATCTGTACCCAGTGGGCT 658

DB 3561 TGTGTTGGGCGCCCGCGTGAACATCATCGCGCGCACATCTGTACCCAGTGGGCT 3620

QY 659 GCACCTGAACTTCCCATCAGCCCGCATCGAGACCGTGCCTGAAGCTGAAGCCCGGCA 718

DB 3621 GCACCTGAACTTCCCATCAGCCCGCATCGAGACCGTGCCTGAAGCTGAAGCCCGGCA 3680

QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778

DB 3681 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGTTGG 3740

QY 779 CCATCTCGGAGGAGATGAGAGGAGGCGCAAGATCAACAAGATCGGCGCGAGAACCCCT 838

DB 3741 AGATCTGCACCGAGATGGAGAAGAGGCGCAAGATCAGCAAGATCGGCGCGAGAACCCCT 3800

QY 839 ACACACCCCGTGTTCGCCATCAAGAGAGGAGACAGCACCAAGTGGCGCGAGCTGTGG 898

DB 3801 ACACACCCCGTGTTCGCCATCAAGAGAGGAGACAGCACCAAGTGGCGCGAGCTGTGG 3860

QY 899 ACTTCCCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGCATCCCCC 958

DB 3861 ACTTCCCGAGCTGAACAAGCGCACCCAGGACTTCTTGGAGGTGAGCTGGGCATCCCCC 3920

QY 959 ACCCGCGCGGCTGAAGAAGAGAGACGTGACCGTCTGGAGTGGGCGACGCTACT 1018

DB 3921 ACCCGCGCGGCTGAAGAAGAGAGACGTGACCGTCTGGAGTGGGCGACGCTACT 3980

QY 1019 TCAGCGTCCCTTGGAGGAGACTTCCGCAAGTACACCGCTTCCACATCCCAAGCATCA 1078

DB 3981 TCAGCGTCCCTTGGAGGAGACTTCCGCAAGTACACCGCTTCCACATCCCAAGCATCA 4040

QY 1079 ACAACGAGAGCCCCCGGATCCGCTACCAAGTACAACGCTGCTGCCAGGCTGGAGGCA 1138

DB 4041 ACAACGAGAGCCCCCGGATCCGCTACCAAGTACAACGCTGCTGCCAGGCTGGAGGCA 4100

QY 1139 GCGCCAGATCTTCCAGAGCAGATGACCAAGATCTTGGAGCCCTTCCGCGCCCGCAAC 1198

DB 4101 GCGCCAGATCTTCCAGTGCAGCATGACCAAGATCTTGGAGCCCTTCCGCGCCCGCAAC 4160

QY 1199 CCGAGATCGTATCTACCA---GGCCCGCTGTAGTGGGAGCGACCTGGAGATCG 1252

DB 4161 CCGAGATCGTATCTACCAAGTACATGGACACCTGTAGTGGGAGCGACCTGGAGATCG 4220

QY 1253 GCCAGACCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGGCTGGGCTTCAACA 1312

DB 4221 GCCAGACCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGCTGGGCTGGGCTTCAACA 4280

QY 1313 CCGCCGACAGAGCAGCAAGAGAGCCCGCTTCTTCTGCCCAT-----CGAGCTGCAC 1366

DB 4281 CCGCCGACAGAGCAGCAAGAGAGCCCGCTTCTTCTGCCCAT-----CGAGCTGCAC 4340

QY 1367 CCGACAAGTGCAGCGTGCAGCCATCGAGCGCCGCGGAGAGGAGCTGGACCGTGAACG 1426

DB 4341 CCGACAAGTGCAGCGTGCAGCCATCGAGCGCCGCGGAGAGGAGCTGGACCGTGAACG 4400

QY 1427 ACATCCAGAACTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 1486

DB 4401 ACATCCAGAACTGGTGGGCAAGCTGAACCTGGGCGCAGCAGATCTACCCCGCATCAAG 4460

QY 1487 TGGCGCAGCTGTGAAGCTGTGTGGCGCGCAAGGCCCTTACCGACATCTGTGCCCTGA 1546

DB 4461 TGGCGCAGCTGTGAAGCTGTGTGGCGCGCAAGGCCCTTACCGAGGTGTGTGCCCTGA 4520

QY 1547 CCGAGAGGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCGCGAGCGCCGTCACG 1606

DB 4521 CCGAGAGGCGGAGCTGGAGCTGGCGGAGAACCGGAGATCTCGCGAGCGCCGTCACG 4580

QY 1607 GCGTGTACTAGACCCCGCAGCAAGGACCTGTGTGGCGGAGATCCAGAAGCAGGCCACGAC 1666

DB 4581 GCGTGTACTAGACCCCGCAGCAAGGACCTGTGTGGCGGAGATCCAGAAGCAGGCCACGAC 4640

QY 1667 AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGACCTGAAGACCGCGCATACGCA 1726

DB 4641 AGTGGACCTACCAAGATCTACCAAGAGCCCTTCAAGACCTGAAGACCGCGCATACGCA 4700

QY 1727 AGATGCGCCGCGCCACACCAACGAGCTGAAGAGCTCACCGAGGCGCTGCAGAAAGATCG 1786

DB 4701 GCATGAAGGGCGCCACACCAACGAGCTGAAGAGCTCACCGAGGCGCTGCAGAAAGATCG 4760

QY 1787 CCATGGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAGG 1846

DB 4761 CCACCGAGAGCATCTGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAAGG 4820

QY 1847 AGACCTGGAGACCTGTGTGAGCCGACTACTTGGAGGCCACCTTGGATCCCCGAGTGGAGT 1906

DB 4821 AGACCTGGAGACCTGTGTGAGCCGACTACTTGGAGGCCACCTTGGATCCCCGAGTGGAGT 4880

QY 1907 TCGTGAACACCCCGCTGTGTGAGCTGTGTACCACTGGAGAGGAGCCCATCATCG 1966

DB 4881 TCGTGAACACCCCGCTGTGTGAGCTGTGTACCACTGGAGAGGAGCCCATCATCG 4940

QY 1967 GCGCCGAGACCTTCTACGTGGAGCGCGCGCCCAACCGGAGACCAAGATCGCAAGGCGC 2026

[illegible]

Search completed: February 10, 2003, 12:45:42
Job time : 582.804 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 11:44:48 : Search time 3377.75 Seconds
(without alignments)
11780.710 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcgacgcaccatggcga.....gggctagcaccggtgaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90.2	3.7	2598	11 AY103647	AY103647 Zea mays
2	85.6	3.5	951	13 BM321451	BM321451 rockefell
3	80.2	3.3	1132	13 BM320864	BM320864 rockefell
4	79.6	3.2	1165	13 BM320900	BM320900 rockefell
5	79	3.2	867	13 BM321430	BM321430 rockefell
6	76.6	3.1	1550	13 BM321022	BM321022 rockefell

7	75	3.1	545	13 BI724851	BI724851
8	75	3.1	862	13 BM321023	BM321023 rockefell
9	73.2	3.0	853	13 BM321393	BM321393 rockefell
10	72.8	3.0	566	13 BM587428	BM587428 170006873
11	72.2	2.9	640	10 BE601575	BE601575 HVSMEH009
12	71.6	2.9	500	13 BM372120	BM372120 EBrO03-SQ
13	71.6	2.9	538	13 BM368580	BM368580 EBrO08-SQ
14	71.6	2.9	540	9 AJ471121	AJ471121 AJ471121
15	71.6	2.9	566	14 BQ464692	BQ464692 HF02P20r
16	71.6	2.9	578	13 BM377112	BM377112 EBrO05-SQ
17	71.6	2.9	579	12 BF253640	BF253640 HVSMEF000
18	71.6	2.9	582	13 BM372871	BM372871 EBrO04-SQ
19	71.6	2.9	588	14 BQ765498	BQ765498 EBrO03-SQ
20	71.6	2.9	600	9 AL508778	AL508778 AL508778
21	71.6	2.9	625	14 BQ768295	BQ768295 EBrO08-SQ
22	71.6	2.9	636	13 BM370222	BM370222 EBrO08-SQ
23	71.6	2.9	660	14 BQ762385	BQ762385 EBrO01-SQ
24	71.6	2.9	700	9 AL508513	AL508513 AL508513
25	71.6	2.9	768	12 BG365317	BG365317 HVSMEI000
26	71.6	2.9	789	12 BF626988	BF626988 HVSMEB000
27	71.6	2.9	822	12 BG417596	BG417596 HVSMEK001
28	71.4	2.9	730	12 BF260696	BF260696 HVSMEF002
29	71.2	2.9	646	13 BM645449	BM645449 170006873
30	71.2	2.9	724	13 BM584191	BM584191 170006872
31	71	2.9	925	17 CNS0091P	AL053013 Drosophil
32	70.6	2.9	861	10 BE636696	BE636696 rockefell
33	70	2.8	494	12 BG349122	BG349122 947029E04
34	70	2.8	571	12 BF266979	BF266979 HV_CEA001
35	70	2.8	609	13 BJ292939	BJ292939 BJ292939
36	70	2.8	613	13 BJ226790	BJ226790 BJ226790
37	70	2.8	616	13 BJ280623	BJ280623 BJ280623
38	70	2.8	619	13 BJ291714	BJ291714 BJ291714
39	70	2.8	627	12 BG300932	BG300932 HVSMEB001
40	70	2.8	635	13 BJ290451	BJ290451 BJ290451
41	70	2.8	643	13 BJ289971	BJ289971 BJ289971
42	70	2.8	652	13 BJ290578	BJ290578 BJ290578
43	70	2.8	659	13 BJ226590	BJ226590 BJ226590
44	70	2.8	660	13 BJ288689	BJ288689 BJ288689
45	70	2.8	662	13 BJ281217	BJ281217 BJ281217

ALIGNMENTS

RESULT 1
AY103647
LOCUS
DEFINITION Zea mays PCO142084 mRNA sequence.
ACCESSION AY103647
VERSION AY103647.1 GI:21206725
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.

linear HTC 25-MAY-2002

2598 bp mRNA

PCO142084 mRNA sequence.

GI:21206725

HTC.

Zea mays.

Zea mays

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2598)

Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,

Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.

Maize Mapping Project/Dupont Consensus Sequences for Design of

Overgo Probes

Unpublished (2002)

2 (bases 1 to 2598)

Coe,E.C.

Direct Submission

Submitted (25-APR-2002) Maize Mapping Project, University of

Missouri, Columbia, MO 65211, USA

Location/Qualifiers

1.2598

/organism="Zea mays"

/db_xref="MaizeDB:638378"

/db_xref="taxon:4577"

/clone="PCO142084"

/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"									
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"									
BASE COUNT	490 a	907 c	795 g	406 t					
ORIGIN									
Query Match	3.7%	Score 90.2;	DB 11;	Length 2598;					
Best Local Similarity	42.2%;	Pred. No. 2e-05;							
Matches 814;	Conservative 0;	Mismatches 1098;	Indels 15;	Gaps					
QY	296	CCACGACGCGGAGCTGCAGGTGGCGGCGACAAACCCCGCAGCGAGGCGCGCGCGAGC	355						
DB	51	CCACCACTCCACCTCTGCCACCGCGGCGGCCCAACACACACCGCAGCGAGCGACAA	110						
QY	356	GCCAGGCAACCCCTGAATTTCCCTCCAGATCACTCTGTGGCAGCCCTCTGTGTGAGCATCA	415						
DB	111	TGGCGGCAACGGCGCATCTGTGAGAGCGACCCGCTGAACCTGGGCGCGCGCGCGG	170						
QY	416	AGTGGGCGGCGAGATCAAGAGGCGCTGTGTGACACCGGCGCGGAGCAGCACCTGCTGTGG	475						
DB	171	AGCTGGCGGAGGACCACTGTGACAGAGGTGAAGCGCATGTGTGGCGCAGGCGCGCAGCCCG	230						
QY	476	AGGAGATGAGCCTTGCCCGGAAGTGAAGGCCAACAGATGATCGGGCGGATCGCGGCTTCA	535						
DB	231	TGCTCAAGATCGAGGTTCCACCTTCGCGTCGCCAGGTGGCGCGCTCGCTTCGCCCA	290						
QY	536	TCAGGTGGCGGAGTACGACAGATCCTGTATCGAGATCTGCGGCAAGAGGCCATCGGCA	595						
DB	291	AGGACGGGTCGCGGTCGCGTTCGAGTCGACGAGGAGGCGCCCGCGCGTCAAGGCCA	350						
QY	596	CGTGTGATCGGCGCCACCCCGTGNACATCATCGGCGCGCAACATGCTGACCCAGCTGG	655						
DB	351	GCAGCGAGTGGATCCTCGACTGTATCGCCACGCGCGGCGACATCTAGGGGTCAACACCG	410						
QY	656	GCTGCAACCTTGAATTTCCCATCAGCCCATCGAGACCGTGGCCGTGAAGCTGAAGCCCG	715						
DB	411	GCTTGGGCGGACCTCCACCGCGGCACCAAGACGGGCGCGCTCCAGTGCAGCTGC	470						
QY	716	GCATGAGCGGCCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGA	775						
DB	471	TCAGGCATCTCAACCGCGGAATCTTCGGCACCGCGCAGCGAGCGGCACACGCTGCCGTCGG	530						
QY	776	CCGCCATCTGGAGGAGATGGAGAAGGAGGCGAAGATCAACCAAGATCGGCCCGGAGAAC	835						
DB	531	AGGTCAACCGCGGCGGATGCTGGTGGCGATCAACACCCCTCTCCAGGGCTACTCCGGCA	590						
QY	836	CCTACAACACCCCGTGTTCGCCATCAAGAAGAGGAGCAGCAACCAAGTGGCGCAAGCTGG	895						
DB	591	TCCGCTTCGAGATCTCGAGGCCATCAGAGCTGTCTCAACCCGGTGTGAGCCCTTGCC	650						
QY	896	TGACTTTCGCGAGCTGAACAAGCCACCCAGGACTTCTGGGAGGTGACGTGGGCATCC	955						
DB	651	TGCCGTTCGGGGCACCATCACCGCTTCGGGCGACCTGGTCCCGCTCTCTACATCCCG	710						
QY	956	CCACCCCGCGGCTGAAGAAGAAAGAGCTGACCGTGTGTGGAGTGGGCGACCGCT	1015						
DB	711	GCCTCATCAGGGCGCGCCCAAGCGCAGGCGCTCACCG---TCGACGGGAGGAAGTGG	767						
QY	1016	ACTTCAAGCTGCCCCGAGGAGCTTCGCGAAGTACACCGCTTCACCATCCCCAGCA	1075						
DB	768	ACGCGCGGAGGCTTCAAGATCCCGGCATCGAGGGCGGCTCTTCAAGCTCAACCCCA	827						
QY	1076	TCACAACAGAGACCCCGGCATCCGCTACCAAGTACACCTGTGTGCGCCAGGGCTGGAAGG	1135						
DB	828	AGGAGGCGCTCGCATCTCAACGGCACGTCGCTGGGCTCGCGCTCGGCGCCACCGTGA	887						
QY	1136	GCAGCCCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCGCCGCA	1195						

DB	888	TGTACGAGCCCAACGTCTCTGGCGGTCTCTGTGGAGGTCCT---GTCCGCGCTCTTCTGGC	944
QY	1196	ACCCCGAGATCGTGTATCTACACGCGCCCTCTGTACGTGGCGAGGACCTGAGATCGGCC	1255
DB	945	AGTTCATGACGCAAGCCCGAGTACAGGACACCTGACCCACAAGCTGAAGACCAACC	1004
QY	1256	AGCACCGCGCAAGATCGAGGAGTGCAGAACACCTGCTG---CGCTGGGGCTTACCA	1312
DB	1005	CGGGTCCATCGAGCGCGCGCCATCATGGAGCACATCTTGGATGGCAGCTTCTTCATGA	1064
QY	1313	CCCCGACAGAAGCACCAGAAGAGCCCTTCTTCCGCCATCGAGCTGCACCCCGACA	1372
DB	1065	AGCAGGCCAAGAGTTGAACAGAGCTGGACCCGCTGCTGAAGCCCAAGCAGGAGTACG	1124
QY	1373	AGTGGACCTGACGCCCATCGAGCTGCCGAGAGGAGAGCTGGACCGTGAACACACATCC	1432
DB	1125	CGTCCGACGTGCCCGCATGGCTGGGCCCCAGATCGAGGTCTATCCGCGCCCGACCA	1184
QY	1433	AGAAGCTGGTGGCAAGCTGAATGGCCAGCCAGATCTATCCCGCGCATCAAGGTGCGCC	1492
DB	1185	AGTCCATCGAGCGCGAGGTCAACTCCGTGAACGA---CAACCCGCTCATCGACGTCCACC	1241
QY	1493	AGCTGTCAAGCTGCTCCGCGCGCCCAAGGCCCTGACCGGACATCGTCCCTGACCCAGG	1552
DB	1242	CGGCAAGCGCTGCACGCGCGCAACTTCCAGGGCAACCCCATCGGCGTGTCTCCATGACA	1301
QY	1553	AGGCGAGCTGGAGCTGGCGGAGAACCGAGATCTCTGCGGAGCGCGTGCACGCGCTGT	1612
DB	1302	ACGCGCGCTCGCCATCGCCCAACATCGGCAAGTCTATGTTGCGGCGATTTCTCCAGTCTG	1361
QY	1613	ACTACGACCCAGCAAGGACCTGGTGGCCGAGATCCAGAAAGAGGCGCACGACAGTGA	1672
DB	1362	TCAACGAGTTCTACAACAACGCGCTCACTCCAACTGGCGGCGAGCGCAACCCAGCC	1421
QY	1673	CTTACAGATCTACCAGGACCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGC	1732
DB	1422	TGGACTACGCTTCAAGGGCACCGAGATCGCCATGGGCTCTCTACTGTCTCGAGCTCCAGT	1481
QY	1733	GCACCGCCACACCAACGAGTGAAGCAGTGCACGAGGCGGTGCAGAAGATCCCATGG	1792
DB	1482	ACCTGGCAACCCCATCACCAACCGTGCAGAGCGGGAGAGCAGCACACCGAGCGTGA	1541
QY	1793	AGAGCAT---CGTGATCTGGGGCAAGACCCCAAGTTTCCGCTTCCCATCCAGAGAGA	1849
DB	1542	ACTCCCTGGGCTCGTCTCGGCCAGGAAGACCGCGAGGCGATCGACATCTGAAGCTCA	1601
QY	1850	CCTGGGAGACTGGTGGACCGACTACTGGCAGGCGCCACTGGATCCCGAGTGGAGTTCG	1909
DB	1602	TGTCGTCCACTATCTGTGGCGCTGTGCCAGGCGCTGSGACCTGCGCCACTCGAGAGA	1661
QY	1910	TGAACACCCCGCTGGTGAAGCTGTGTACAGCTGGAGAGAGGAGGCCCATCATCGCG	1969
DB	1662	ACATCAAGGCTCGTGAAGAACACCGTGAACCCAGGTGGCCAGAGAGTCTGACCATGA	1721
QY	1970	CCGAGACTTCTACGTGGAGCGGCCCGCCCAACCGCAGACCAAGATCGCAAGCGCGCT	2029
DB	1722	ACCCCTCGGCGAGCTCTCCAGCGCGCGCTTACGAGAGAGGAGCTGATCAGCGCCATCG	1781
QY	2030	ACGTGACCGAGCGGGCGCGCAAGATCTGTAGCCCTGACCCGAGACCAACCAAGAGA	2089
DB	1782	ACGCGAGGCGGTTCACGTACGCGGAGGACGCGGCGAGCGCTGCCGCTGTATGC	1841
QY	2090	CCGAGTGCAGGCCATCTCAGCTGCCCTGCAGGACAGCGGAGGAGGTGAACATCTGA	2149
DB	1842	AGAAGTGGCGCGCTGTGTGTGACCAACCGCTTCAACGCGGCGAGCGGAGCGGAGC	1901
QY	2150	CCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCAGACCAAGAGGAGAGCGAGC	2209
DB	1902	CCTCCGTGTTCTCCAAGATCACCAGGTTTCGAGGAGGAGCTCCGCGCGGTGCTGCCCCAGG	1961
QY	2210	TGTTGA 2216	
DB	1962	AGTTGA 1968	

```
RESULT 2
BM321451
LOCUS
DEFINITION
  639 GTTCCGCGCGTCAAGTCAACGAC---TCTNTACCAAGAGCAAGTTCGACAACATCTA 695
  2202 GAGCGAGCTGGTGAACCAAGATCATCGAGCAGCTGTATCAAGAAGAGAGAGTGTACTCTGAG 2261
  696 CGGCTGCGCGCCACTCGCTCATCGAGGGATCAAGCGCGGCACCGCAGCTGATGCTCGGCGG 755
  2262 CTGGGTGCGCGCCCAAGAGGATCGGCGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCAA 2321
  756 CAAGTTCGCGCTGCTGCGGCGCTACGGCAGCTGGGCAAGGCTGCGCGAGTCTGCTGCG 815
  2322 GGGCATCCCAAGGTGCTTCTCTGGAGCGCATCGATGGCGGCATCGTGTATACAGTA 2381
  816 CGGCGAGGCTGCGCGCTCATGTCGACGAGATCGACCCCATCTGCGGCTCGAGCGCTC 875
  2382 CATGACACCACTGTAGTGGGCGGCGGCGGCGCTAGGATCGA 2423
  876 GATGCGCGGCTTCGAGGTCAACACACGCTCGAGGCGGGGCTCGA 917

ACCESSION
  BM320864
VERSION
  BM320864.1
KEYWORDS
  EST.
SOURCE
  Mastigamoeba balamuthi.
  Mastigamoeba balamuthi.
  Mastigamoebidae; Mastigamoebae.
REFERENCE
  1 (bases 1 to 1132)
  Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
AUTHORS
  Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
  Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
  Philippe,H.
  The analysis of 100 genes supports the grouping of three highly
  divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  21819461
  Contact: Muller Miklos
  Laboratory of Biochemical Parasitology
  The Rockefeller University
  1230 York Avenue, New York, NY 10021, USA
  Email: mmuller@rockvax.rockefeller.edu
  Insert Length: 1132 Std Error: 0.00
  POLYA-No.

JOURNAL
  MEDLINE
  COMMENT

FEATURES
  source
  1..1132
  /organism="Mastigamoeba balamuthi"
  /strain="ATCC 30984"
  /db_xref="taxon:108607"
  /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
  /note="syn: Phreatamoeba balamuthi"
  BASE COUNT 220 a 413 c 324 g 171 t 4 others
  ORIGIN

Query Match 3.3%; Score 80.2; DB 13; Length 1132;
Best Local Similarity 45.3%; Pred. No. 0.00085;
Matches 375; Conservative 0; Mismatches 444; Indels 9; Gaps 2;

  139 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 198
  79 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 138
  199 TGCACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 258
  139 TACAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 192
  259 CGCGAGTTCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 318
  193 TGCAGATCGCGCTACGCCAAGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 252
  319 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 378

RESULT 3
BM320864
LOCUS
DEFINITION
  639 GTTCCGCGCGTCAAGTCAACGAC---TCTNTACCAAGAGCAAGTTCGACAACATCTA 695
  2202 GAGCGAGCTGGTGAACCAAGATCATCGAGCAGCTGTATCAAGAAGAGAGAGTGTACTCTGAG 2261
  696 CGGCTGCGCGCCACTCGCTCATCGAGGGATCAAGCGCGGCACCGCAGCTGATGCTCGGCGG 755
  2262 CTGGGTGCGCGCCCAAGAGGATCGGCGGCAACGAGCAGCAGATCGACAAGCTGGTGAGCAA 2321
  756 CAAGTTCGCGCTGCTGCGGCGCTACGGCAGCTGGGCAAGGCTGCGCGAGTCTGCTGCG 815
  2322 GGGCATCCCAAGGTGCTTCTCTGGAGCGCATCGATGGCGGCATCGTGTATACAGTA 2381
  816 CGGCGAGGCTGCGCGCTCATGTCGACGAGATCGACCCCATCTGCGGCTCGAGCGCTC 875
  2382 CATGACACCACTGTAGTGGGCGGCGGCGGCGCTAGGATCGA 2423
  876 GATGCGCGGCTTCGAGGTCAACACACGCTCGAGGCGGGGCTCGA 917

ACCESSION
  BM320864
VERSION
  BM320864.1
KEYWORDS
  EST.
SOURCE
  Mastigamoeba balamuthi.
  Mastigamoeba balamuthi.
  Mastigamoebidae; Mastigamoebae.
REFERENCE
  1 (bases 1 to 1132)
  Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.
AUTHORS
  Baptiste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
  Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
  Philippe,H.
  The analysis of 100 genes supports the grouping of three highly
  divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
  Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
  21819461
  Contact: Muller Miklos
  Laboratory of Biochemical Parasitology
  The Rockefeller University
  1230 York Avenue, New York, NY 10021, USA
  Email: mmuller@rockvax.rockefeller.edu
  Insert Length: 951 Std Error: 0.00
  POLYA-No.

JOURNAL
  MEDLINE
  COMMENT

FEATURES
  source
  1..951
  /organism="Mastigamoeba balamuthi"
  /strain="ATCC 30984"
  /db_xref="taxon:108607"
  /clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"
  /note="syn: Phreatamoeba balamuthi"
  BASE COUNT 186 a 321 c 303 g 139 t 2 others
  ORIGIN

Query Match 3.5%; Score 85.6; DB 13; Length 951;
Best Local Similarity 46.0%; Pred. No. 0.00011;
Matches 323; Conservative 0; Mismatches 376; Indels 3; Gaps 1;

  1722 CGCCAAAGATGGGACCGCCACACACAGCAGCTGAAGCAGCTGACCGAGCGCGTGCAGAA 1781
  219 CGCCAAAGATGGGACCGCGTGTCTGCTGCAACATCTTCTGACGCGCAGCAGCGCGCGCGC 278
  1782 GATCGCATGAGAGCATGTCGTGGGCAAGACACCCCAAGTTCGCTGCGCTGCCATCCA 1841
  279 CATCGCGAGCGCGCGTCTGCGGCTTGGGAGGCGGAGAACCTCCAGGAGTACTG 338
  1842 GAAGGAGACCTGGGAGACCTGGTGGACCGACTACTGCGAGCGCCACCTGGATCCCGAGTG 1901
  339 GGATGACCTGGAGCGCCCTGTCTGCGCCCTACAGGCGCCCTCAGATCATCTGTCGA 398
  1902 GGAGTTCGTGAACACCGCCCGCTGGTGAAGCTGTGTACCTGAGTGGAGAGGAGCGCAT 1961
  399 CGAGCGCGGTGACGCGACTCTGATGATCAACAGGGGTTTCGCGCGCGAGGACACCCCAA 458
  1962 CATCGGCGCGGAGACCTTCTACGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2021
  459 GCTGCTGGAGGACGACGAGGCGCTTCGAGGAGGTGCGCTCAACACGCTGCTCAAGCA 518
  2022 GGGCGGCTAGCTGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2081
  519 GTTCCAGAGGAGCGCGCGCTTCTGGACACAGATCTCTCCCGAGATCCGCGGTGTGAG 578
  2082 CCAGAGACCGAGCTGACGCGCATTCAGCTGGCGCTTCGAGGACGAGCGCGCGCGCGCGG 2141
  579 CGAGGAGACGACGACTGGCGCTNATGAGGCTGTACCACTGACCGCGCGCGCGCGCGCGCT 638
  2142 CATCGTACCGGACGAGCAGTACGCGCTTGGGATCATCCAGCGCGCGCGCGCGCGCGCGCG 2201
```

```
Db 253 GAGCTACCCGCTGCGGCTCAAGCTCGGCTGACCAACTACGCGCGGCTACGCGACT 312
Qy 379 CAGATACCCCTGTGGCAGCGCCCTGGTGAGCATCAAGTGGCGGCCAGATCAAGAG 438
Db 313 GGCTGTGCTGCGGCCCGCTGTGTAAGAAGCTCAACCTCGACTCCAAGTAGAGGGT 372
Qy 439 GCCTGTGGACACCGCGCGGACACACGCTGTGGAGGAGATGAGCCTGCCCGGCAAG 498
Db 373 GTCAAGAAGTCAACGCGGAGGACTACAAGTCAGAGAGTCGACGAGCGGCCCGCT 432
Qy 499 TGAAGCCCAAGATGATGCGCGGATCGGCGGCTTTCATCAAGTGGCGCCAGTAGACAG 558
Db 433 TTCAAGCCCTGTGCGACGTCGGCGCTGGTCCGCAACCTCGACTGGCGCGCGCTTCGCC 492
Qy 559 ATCCTGATGAGATCTCGCGCAAGAAGGCAATCGGACCGCTGTGATGCGGCCACCC 618
Db 493 GCCCTCAAGGGCATGTGCGACGGGGGCTCAACGTCCCC---CACAGCGAGACCGGCTTC 549
Qy 619 GTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCTGCACCTGAACTTCCCCATC 678
Db 550 GTGGCTTCAACGGGACAGAGAGCTCAAGCCGCGGTCTCTCCGCAAGTACATCTTC 609
Qy 679 AGCCCCATGAGACCGTGGCGGTGAAGCCCGGCGATGAGCGGCCCAAGGTGAAG 738
Db 610 GCGCGCACGTCGCGGCTACATGAAGCTCCTCAAGGAGCAGGACGCGCGCTTCGAC 669
Qy 739 CAGTGGCCCTGACCGGAGGAGATCAAGGCCCTGACCGCCATCTCGGAGGATGGAG 798
Db 670 CGCCAGTTCGCGCTAGCCAGAGAGGTGTACCGCGGACATGCTCGAGAAGATCTAC 729
Qy 799 AAGGAGGCAAGATCAACAGATCGGCGCCGAGAACCCCTACAAACCCCGTGTGCGCC 858
Db 730 ACCGAGCCCAACAGAGATCGCGCGGACCGGACCTTCGTCCCAAGCGCGCTCGAAG 789
Qy 859 ATCAAGAAGAAGGACAGACCAAGTGGCGCAAGCTGGTGGACTTCGCGGAGCTGAACAAG 918
Db 790 CCGGAGGCGCAAGCCCAAGCACTGGGCGAAGCGGAGGTGAGCTACCAAGCGGCAAG 849
Qy 919 CGCACCCAGGACTTCTGGAGGTGACGTGGGCATCCCGACCCCGCC 966
Db 850 AACCGCTGCGCCAGAGAGGTCGCGCTGGGCTACCCCGACGCCCC 897
```

RESULT 4
BM320900 1165 bp mRNA linear EST 03-JAN-2002
LOCUS rockefeller.0.353 Mastigamoeba balamuthi lambda ZAP II Library
DEFINITION Mastigamoeba balamuthi cDNA similar to ribosomal protein L5, mRNA
sequence.
BM320900
BM320900.1 GI:18055306
EST
SOURCE Mastigamoeba balamuthi.
ORGANISM Mastigamoeba balamuthi
Eukaryota; Pelobiontida; Mastigamoebidae; Mastigamoeba.

REFERENCE 1 (bases 1 to 1165)
Bapteste,E., Brinkmann,H., Lee,J.A., Moore,D.V., Sensen,C.W.,
Gordon,P., Durufle,L., Gaasterland,T., Lopez,P., Muller,M. and
Philippe,H.
The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
21819461
Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1165 Std Error: 0.00
POLYA-No.

TITLE The analysis of 100 genes supports the grouping of three highly
divergent amoebae: Dictyostelium, Entamoeba, and Mastigamoeba
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (3), 1414-1419 (2002)
MEDLINE 21819461
COMMENT Contact: Muller Miklos
Laboratory of Biochemical Parasitology
The Rockefeller University
1230 York Avenue, New York, NY 10021, USA
Email: mmuller@rockvax.rockefeller.edu
Insert Length: 1165 Std Error: 0.00
POLYA-No.

FEATURES
source Location/Qualifiers
1..1165

```
/organism="Mastigamoeba balamuthi"  
/strain="ATCC 30984"  
/db_xref="taxon:108607"  
/clone_lib="Mastigamoeba balamuthi lambda ZAP II Library"  
/note="syn: Phreatamoeba balamuthi"  
BASE COUNT 221 a 426 c 336 g 178 t 4 others  
ORIGIN  
Query Match 3.2%; Score 79.6; DB 13; Length 1165;  
Best Local Similarity 45.1%; Pred. No. 0.0011;  
Matches 428; Conservative 0; Mismatches 506; Indels 16; Gaps 3;  
Qy 90 CATCAAGTGTTCACATCGCGCAAGGAGGCGCCACATCGCGCGCAACTGCGCGCCCGCCG 149  
Db 46 CGTCAAGAACAAAGCGTACTTCAAGCGCTTCCAGACCCAGTTCCGTGCGCGCGGAGG 105  
Qy 150 CAAGAAGGGTGTGGAAGTGGCGCAAGGAGGCGCCACCAGATGAAGGACTGCACCGAGCG 209  
Db 106 CAAGACGAGCTTACCGCGCGGCCANCTGGTGATCCAGGACAAAGTACAACAGCCC 165  
Qy 210 CCAGGCCAACTTCTTCGCGAGGACCTGGCCTTCCCGCAGGCAAGGCCCGCGAGTTCCC 269  
Db 166 CAAGTACCGCTTCGTC-----GTCCGCTTCAACCAACAGGACATCGTCTGCCAGATCGC 219  
Qy 270 CAGCGAGCAAGACCGCGCAACAGCCCGCACAGCGCGGAGCTGCAGGTGCGCGGAGCAA 329  
Db 220 CTACGCGAAGATCACGCGGACACATCTCGCGCGCCCTACTCGCACGAGCTCACCCG 279  
Qy 330 CCCCCGAGCAGCGCGCGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 389  
Db 280 CTTGGCGGTCAAGTTCGGCTTGACCAACTACGCGCGCGCTTACCGGACTGGCGCTGCTGCT 339  
Qy 390 GTGCAGCGCGCCCTGTGTGAGCATCAAGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 449  
Db 340 GCGCGCGCTGTGTGAAGAGCTCAACCTCTACTCCAAGTACGAGGCTGCAAGAAGGT 399  
Qy 450 CACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 509  
Db 400 CAAGGCGGAGGACTACACGCTCGAGGAGCTCGAGGAGCGCGCGCGCGCGCGCGCGCGCG 457  
Qy 510 GATGATGCGCGGCGATCGCGCGCTTATCAAGGTGCGCGGCGGCGGCGGCGGCGGCGGCGG 569  
Db 458 --TGCTGACGCTCGGCTGTGTCGCGACCTCGACTGGCGCGCGCGGCTGTTCGCGCGCGCTCAA 515  
Qy 570 GATCTGCGCGAAGAGGCGCATCGGCGCGCTGTGATCGCGCGCGCGCGCGCGCGCGCGCGCAT 629  
Db 516 GGGCATGTGCGCGCGCGCGCTCAAGTCCCGCACAGCGAGAGCGCGCTTCGTGGGCTTCAA 575  
Qy 630 CGCGCGCAACATGTGTGACCGAGCTGGGCTGCACCTGCACTTCCCGATCAGCGCGCGCGCG 689  
Db 576 CGGCGCAAGAAGGAGGCTCAACGCGCGGCTCTCGGCAAGTACATCTTCGCGCGCGCGCG 635  
Qy 690 GACGTCGCGCTGAAGCTGAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 749  
Db 636 CGCGCGGTACATGAAGCTCTCAAGGACGAGACGCGCGCGCGCTTCGACCGCGCGCTTCTC 695  
Qy 750 GACCGAGGAGAGATCAAGGCGCTGACCGCGCTCTGCGGAGGAGATGGAGAAGGAGGCGAA 809  
Db 696 CGGTACGCCAAGGAGGAGGTGTACCGCGGACATGTCTGAGAAGATCTACACGAGGCGCCA 755  
Qy 810 GATCACCAAGATCGGCGCGCGGAGACCCCTCAACACCCCGCTGTTCGCCATCAAGAAGAA 869  
Db 756 -----CAAGCAGATTCGCGCGCGGACCGCGCTCTGTCGCGCGCGCGCTCGAAGCGCGA 809  
Qy 870 GGACGACCAAGTGGCGCGCAAGCTGGTGGAGCTTCGCGGAGCTGGAACAACAGCGCACCGA 929  
Db 810 GGGCGCCCAAGCCCAAGCACTGGGCGCAAGCGGAGGCTGAGTACCAAGCGGCGCAAGACCG 869  
Qy 930 CTTCTGGGAGGTGCGAGCTGGCGCATCCCCACCGCGCGCGCTTGAAGAAGAAGAGCGGT 989  
Db 870 CGTCGCCCAAGAAGAGGTCCCGCTGGGCTACCCCGACGCGCGCGCGCGCGCGCGCGCGCG 929  
Qy 990 GACCGTGTGGAGTGGCGGAGCGGCGCTACTTCAGCGTGGCGCGCGCGCGCGCGCGCGCG 1039
```

15 JULY 2005


```

Db 596 GAGGGCCCAAGCCCAAGCACTGGGGCAAGCGCAGGCTGACGTACCAGCAGCCGAAGAAC 655
QY 928 GACTCTCTGGGAGGTGTCAGCTGGGCGATCCCCACCCCGCCCTGAAGAAGAAAGAACAGC 987
Db 656 CGCGTCCCAAGAGAGGTCCGCTGGGTACCCCGACGCCCCCAAGACCGAGCAATTT 715
QY 988 GTGACCTGTGTGACGTGGGCGACGCTACTTTCAGCG 1024
Db 716 CCGGTGCCCTGTACACGCGCATCTCCCGCTGGCGG 752

RESULT 10
BM587428/c
LOCUS
DEFINITION
  BM587428 566 bp mRNA linear EST 25-FEB-2002
  19600449696110 5', mRNA sequence.
ACCESSION
  BM587428
VERSION
  BM587428.1 GI:18883289
KEYWORDS
  African malaria mosquito.
SOURCE
  Anopheles gambiae
ORGANISM
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
  Anopheles.
REFERENCE
  1 (bases 1 to 566)
  Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
  ,R., Collins,F.H., Venter,J.C. and Hoffman,S.L.
  Celera Anopheles gambiae EST project
  Unpublished (2002)
  Contact: Holt R.A.
  Celera Genomics
  45 W. Gude Dr., Rockville, MD 20850, USA
  Tel: 2404533151
  Fax: 2404534580
  Email: Holt@raecelera.com
  Plate: NU01004AAX row: G column: 08
  Seq primer: M13 Reverse.
FEATURES
  source
  1..566
    /organism="Anopheles gambiae"
    /strain="RSP-ST (Reduced susc. to Permethrin - std.
    chromosome)"
    /db_xref="taxon:7165"
    /clone="19600449696110"
    /clone_lib="A.Gam.ad.cDNA.bloodi"
    /dev_stage="Adult"
    /lab_host="DH10b"
    /note="Vector: pSport1; Site_1: SalI; Site_2: NotI; Whole
    adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
    hours after human blood feeding. cDNA inserts >500 bp
    cloned directionally into pSport 1. Not 1 site is 3'.
    Clones available through the Malaria Research and
    Reference Reagent Resource Center (www.malaria.mr4.org)"
  BASE COUNT 60 a 202 c 165 g 139 t
  ORIGIN
    Query Match 3.0%; Score 72.8; DB 13; Length 566;
    Best Local Similarity 48.3%; Pred. No. 0.014;
    Matches 203; Conservative 0; Mismatches 217; Indels 0; Gaps 0;

QY 1944 GTTGGAGAGGAGCCCATCATCGCGCGAGACATCTTCTAGTGGACGGCGCCCAACCG 2003
Db 482 GTTCGAGATCGACGCCAAGCGCATCTGCGGTGTCGGCGGAGCAAGGCGCGGCA 423
QY 2004 CGAGACCAAGATCGGAAGCCCGCTACGTACCGACCGGGCGCGAGAGATCGTGAG 2063
Db 422 CCGGAAAAGATCGTCATCACCACCAAGCAGAACCCGCTGACCCCGGAGACATCGAGCG 363
QY 2064 CTGTACCGGAGACCAACCAAGACCGAGCTGACGAGGCATCCAGCTGGCCCTGCAGGA 2123
Db 362 CATGATCAAGGATCGGAGCGGTTCGCCCGACGACGACACAGAGCTGAAGGAGCGCGTGA 303
QY 2124 CACGGCAGGAGGTGAATCTGTGACCGACACCCAGTACGCGCTGGGCATCATCCAGGC 2183

```

```

Db 302 GGCCCGCAACGAGCTCAGAGCTACGCGTGAAGAACCCAGCTCAGCTCGAAGGA 243
QY 2184 CCAGCCCGCAAGAGCGCAGAGCGAGCTGGTGAACCAACCATATCGCAGCAGCTGATCAAGAA 2243
Db 242 CAAGCTGGGCGGAGCGGTGTCGCGACGACGACAAAGGCCAAGATGGAGGCGGATCGACGA 183
QY 2244 GGAGAAGGTGTACCTGAGCTGGGTGCCCGCCCAAGGGCATCGCGGCAACGAGCAGAT 2303
Db 182 GAAGATCAAGTGGCTGACGAGAACACGAGGACACCGAGCGGGAAGAGTACAAGAAGCAGAA 123
QY 2304 CGACAAGCTGGTGAGCAAGGCGCATCCGCAAGGTGCTGTTCCTGCGACGGCGATCGATGCGG 2363
Db 122 GAAGGAGCTGGAGGACATCGTGACAGCCCATCATTTGCCAAGCTGTACGCGAGCAGTGGCGG 63

RESULT 11
BE601575
LOCUS
DEFINITION
  BE601575 640 bp mRNA linear EST 22-OCT-2001
  HVSMEH0098K08f Hordeum vulgare 5-45 DAP spike EST library
  HVCDNA0009 (5 to 45 DAP) Hordeum vulgare cDNA clone HVSMEH0098K08f,
  mRNA sequence.
ACCESSION
  BE601575
VERSION
  BE601575.3 GI:16322423
KEYWORDS
  Hordeum vulgare.
SOURCE
  Hordeum vulgare
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Hordeum.
  1 (bases 1 to 640)
REFERENCE
  Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
  ,Y., Henry,D., Palmer,M., Rambo,M., Simmons,J., Choi,D.W., Fenton
  ,R.D., Close,S.J., Oates,R. and Main,D.
  Development of a genetically and physically anchored EST resource
  for barley genomics: Morex 5-45 DAP spike cDNA library
  Unpublished (2001)
  On Aug 21, 2000 this sequence version replaced gi:13190104.
  Contact: Wing RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson, SC 29634, USA
  Tel: 864 656 7288
  Fax: 864 656 4293
  Email: rwing@clemson.edu
  Total hg bases = 461
  Seq primer: AATTAACCTCTACTAAAGG
  High quality sequence stop: 534.
FEATURES
  source
  1..640
    /organism="Hordeum vulgare"
    /cultivar="Morex"
    /db_xref="taxon:4513"
    /clone="HVSMEH0098K08f"
    /clone_lib="Hordeum vulgare 5-45 DAP"
    /tissue_type="5-45 DAP Spike"
    /lab_host="SOLR"
    /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI;
    Plants were grown in the greenhouse at the University of
    California, Riverside (Fenton, SJ Close, TJ Close). Whole
    spikes with awns trimmed were collected at 5, 10, 15, 20,
    30 and 45 DAP (Fenton). Total RNA was prepared from each
    pool, equal quantities of all six RNA pools were combined,
    poly(A) RNA was purified from the mixture, one primary
    unamplified cDNA library was made, and 1 million pfu were
    in vivo excised to give pluescript SK(-) cDNA phagemids
    (Choi) in the TJ Close lab at the University of California,
    Riverside. Phagemids were plated and picked at the Clemson
    University Genomics Institute (CUGI) (Begum, Palmer,
    Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
    sequencing and sequence analysis were performed at CUGI
    (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).

```

The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see

<http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinjohs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/gppages/bgn/31/cover.html>)

```

BASE COUNT      135 a      230 c      177 g      97 t
ORIGIN

Query Match      2.9%; Score 72.2; DB 10; Length 640;
Best Local Similarity 47.7%; Pred. No. 0.017;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAAGAGGGCAAGATCACAGATCGGCGCGAGAACCCCTACAAACACCC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 110 GGAGAGGTGACGGTGAAGGTGTCGGCAAGATGATCTCGGTGACGGCGCGGCAC 169

QY 849 CGTGTTCGCCATCAAGAGAGGAGCACCACCAAGTGGCGCAAGCTGTGGACTTCCGCGA 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 170 CTTGACCGCAACTTCAAGCACCTCAACTCGACTTCCAGCTGCAGGACGGCGGCACAA 229

QY 909 GCTGAACAAGCGCACCGAGGACTTCTGGAGGTGCAGCTGGGCATCCCGCACCCCGCGG 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 GCTCAAGGTGACCGCTGTTCGCGACCCCGCGCACCATGCGCGCATCCGCCACCGCAT 289

QY 969 CTTGAAGAAGAAGAGCGTGACCGGTGCTGGAGGTGGCGGACCGCTTACTTCAGCGTCC 1028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 CTCCACGTCAGAACCTCATACCGCGGTACCAAGGGCTTCGGCTTACAAGATGCGCTT 349

QY 1029 CTTGAGAGGACTTCCGCAAGTACCGGCTTCACCATCCCGAGATCAACACAGAGAC 1088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 350 CGTCTAGCTCACTTCCCGCATACAGCGCTCATACCGCGCGCAACCGCGGCATCGAGAT 409

QY 1089 CCGCGGATCCGTACCAAGTACAGCTGCTGCCCGCGGCTTGAAGGCGACCCCGCAT 1148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 CCGCAACTTCTTCGCGGAGAGAGGTGAGGAGGTGACATGCTCAGCGGGTCAACCAT 469

QY 1149 CTTCCAGAGCATGACCAAGATCTTGGAGCCCTTCGCGCGCGCAACCCCGAGATCGT 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 CTTCCGCTCGGAGAGGTCAAGATGATGATCGTCTCTCGAGGCAACGACATCGAGCTCGT 529

QY 1209 GATCTACAGGCCCGCCCT 1226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 NTCCGCTCCGCGCCCT 547

```

```

RESULT 12
BM372120
LOCUS
DEFINITION
  Ebro03_SQ004_B09_R root, 3 week, waterlogged, cv Optic, Ebro03
  Hordeum vulgare cDNA clone Ebro03_SQ004_B09 5', mRNA sequence.
ACCESSION
  BM372120
VERSION
  BM372120.2 GI:21948499
KEYWORDS
  EST.
SOURCE
  Hordeum vulgare.
  Hordeum vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
  ; Triticeae; Hordeum.

```

```

REFERENCE
  1 (bases 1 to 500)
  Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
  Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  Development of Barley Transcriptome Resources
  Unpublished (2001)
  On Jan 10, 2002 this sequence version replaced gi:18115510.
  Contact: Waugh R, Marshall DF
  Genome Dynamics/Computational Biology
  Scottish Crop Research Institute
  Invergowrie, Dundee, DD2 5DA, Scotland, UK

```

```

Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: est@scri.sari.ac.uk
All sequence has a Phred quality score of 20 or over
Seq primer: M13 reverse.
Location/Qualifiers
  source
    1..500
    /organism="Hordeum vulgare"
    /cultivar="Optic"
    /db_xref="taxon:4513"
    /clone="Ebro03_SQ004_B09"
    /clone_lib="root, 3 week, waterlogged, cv Optic, Ebro03"
    /tissue_type="root"
    /dev_stage="3 week"
    /lab_host="DH10B"
    /note="vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
    Non-normalised library, directionally cloned into pSPORT1.
    Derived from roots of 3 week old waterlogged barley
    plants. Developed as part of the barley transcriptome
    resources of BBSRC/SEERAD funded cereal IGF (Investigating
    Gene Function) project."
  BASE COUNT      108 a      177 c      139 g      76 t
ORIGIN

```

FEATURES

```

Query Match      2.9%; Score 71.6; DB 13; Length 500;
Best Local Similarity 47.7%; Pred. No. 0.021;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 789 GGAGATGGAAGAGGGCAAGATCACCAAGATCGGCGCGAGAACCCCTACAAACACCC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 32 GGAGAGGTGACGGTGAAGGTGTCGGCAAGATGATCTCGGTGACGGCGCGCGGCAC 91

QY 849 CGTGTTCGCCATCAAGAGAGGAGCACCACCAAGTGGCGCAAGCTGTGGACTTCCGCGA 908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 92 CTTACCGCGCAACTTCAAGCACCTCAACCTCGACTTCCAGCTGCAGGACGGCGGCACAA 151

QY 909 GCTGAACAAGCGCACCCAGGAGTCTTGGGAGGTGCAGCTGGGCATCCCGCACCCCGCGG 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 152 GCTCAAGGTGACCGCTGTTCGCGACCCCGCGCACCATGCGCGCATCCGCGCGCAT 211

QY 969 CTTGAAGAAGAAGAGCGGTGACCGTCTGGAGTGGCGGACCGCTTACTTCAGCGTCC 1028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 212 CTCCACGCTCCAGAACCTCATCACCGCGGTCCACCAAGGGCTTCCGCTACAAGATCGCGTT 271

QY 1029 CTTGAGAGGACTTCCGCAAGTACACCGCTTCCACATCCCGAGATCAACACAGGAC 1088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 272 CGTCTAGCTCACTTCCCGCATCAACCGCTTCATCACCGCGCGCAACCGCGGCATCGAGAT 331

QY 1089 CCGCGGATCCGCTACCAAGTACAACTGCTGCCCGCGGCTTGAAGGCGACCCCGCAT 1148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 332 CCGCAACTTCTTCGCGGAGAGAGGTGAGGAGGTGACATGCTCAGCGGGTCAACCAT 391

QY 1149 CTTCCAGAGCATGACCAAGATCTTGGAGCCCTTCGCGCGCGCAACCCCGAGATCGT 1208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 392 CTTCCGCTCGGAGAGGTCAAGGATGATGATCGTCTCTCGAGGCAACGACATCGAGCTCGT 451

QY 1209 GATCTACAGGCCCGCCCT 1226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 CTCCCGCTCCGCGCCCT 469

RESULT 13
BM368580
LOCUS
DEFINITION
  Ebro03_SQ004_A06_R embryo, 40 DPA, no treatment, cv Optic, Ebro08
  Hordeum vulgare cDNA clone Ebro08_SQ004_A06 5', mRNA sequence.
ACCESSION
  BM368580
VERSION
  BM368580.2 GI:21937722
KEYWORDS
  EST.
SOURCE
  Hordeum vulgare.
  Hordeum vulgare
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

```


LOCUS B0464692 566 bp mRNA linear EST 30-MAY-2002
DEFINITION HF02P20r HF Hordeum vulgare cDNA clone HF02P20 5-PRIME, mRNA
sequence.
ACCESSION B0464692
VERSION B0464692.1 GI:21272474
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare
REFERENCE
AUTHORS Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.
TITLE Barley ESTs from developing seeds
JOURNAL Unpublished (2002)
COMMENT Contact: Stejneger Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stejn@ipk-gatersleben.de
Insert Length: 566 Std Error: 0.00
Plate: 2 row: p. column: 20
Seq primer: M13rev.

FEATURES
source
1..566
Location/Qualifiers
/organism="Hordeum vulgare"
/cultivar="Barke"
/db_xref="taxon:4513"
/clone="HF02P20"
/clone_lib="HF"
/tissue_type="caryopsis"
/dev_stage="developing caryopsis, 16-25 DAF (days after
flowering)"
/lab_host="XL10-Gold"
/note="vector: pBluescript SK⁺; Site_1: EcoRI (5'-end of
cDNA); Site_2: XhoI (3'-end of cDNA); developing caryopsis
, 16-25 DAF(days after flowering) Due to a cloning artefact
caused by the kit, in most cases the EcoRI site is NOT
present, as well as the EcoRI adapter used for cloning. To
excise the insert, restriction sites upstream EcoRI should
be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the
cloning system used Blue/white selection for recombinants
is not 100% reliable. Average insert size is 940 bp"

BASE COUNT
ORIGIN

Query Match 2.9%; Score 71.6; DB 14; Length 566;
Best Local Similarity 47.7%; Pred. No. 0.021;
Matches 209; Conservative 0; Mismatches 229; Indels 0; Gaps 0;
QY 789 GGAGATGGAGAGGAGGCAAGATCACCAAGATCGGCCCGCCAGAACCCCTACACACACCC 848
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GGAGAGGTGACGGTGAAGGTGTGCGGCAAGATGATCTCGGTGACGGGGCGCGCGGCAC 61
QY 849 CGTGTTCGCCATCAAGAGAGGAGGACACCAAGTGGCGCAAGCTGGTGGACTTCCGCGA 908
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 CCTGACCCGCAACTTCAAGCACCTCAACCTCGACTTCCAGCTGACGAGCGCGGGCGCAA 121
QY 909 GCTGAACAGCGGACCCAGACACTTCTGGAGGTGCGAGTGGGATGCCACCCACCCCGCGG 968
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 GCTCAAGGTGGAGCGCTGGTTCGGSCACCGCGCACCCATGGCGCCATCGGCACCGCCAT 181
QY 969 CCTGAAGAAGAAGAGCGTGCCTGGAGTGGCGGACGCTTACTTCAGCGTGCC 1028
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 CTCCCAGCTCAGAACCTATACCGCGGTACCAAGGGCTTCGGCTACAGATGCGCTT 241
QY 1029 CTTGGAGGAGGACTTCGGCAAGTACACCGCTTACCATCCCGAGCATCAACACAGAGAC 1088
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 CGTCTAGGCTCACTTCCCATCAAGGCTTCCATCACCAGCCGCAACCGGGGATCGAGAT 301
QY 1089 CCCCAGCATCCGTACCAAGTACAGTGTGCCCCAGGGCTGGAAGGGGACCCCGCAT 1148

Db 302 CGCAACTTCTTCGCGGAGAGAGGTGAGGAGGTGGACATGCTCGAGGGGTCACCAT 361
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1149 CTTCCAGAGCAGCATGACCAAGATCCTTGAGGCCCTTCCGCGCCCGCAACCCCGAGATCGT 1208
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 CTTGCGGTCCGAGAGGTCAAGGATGAGATCGTCTCGACGGCAACGACATCGAGTCTCT 421
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1209 GATCTACCAAGGCCGCCCT 1226
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 CTCCGCTCCGCCGCCCT 439
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: February 10, 2003, 20:43:52
Job time : 3389.75 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:14:58 ; Search time 79.8051 Seconds
(without alignments)
9441.811 Million cell updates/sec

Title: US-09-610-313-32
Perfect score: 2457
Sequence: 1 gtcagccaccatggccga.....gggctagcaccgtgaattc 2457

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15333831 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081.6	44.0	2601	4	US-09-117-217-7
2	1081.6	44.0	2601	4	US-09-117-217-9
3	1081.6	44.0	2601	4	US-09-117-217-11
4	1081.6	44.0	2601	4	US-09-117-217-13
5	1075.2	43.8	7399	2	US-08-418-848A-9
6	1075.2	43.8	9709	2	US-08-188-583-5
7	1075.2	43.8	9709	3	US-08-388-353-1
8	1075.2	43.8	9709	3	US-08-488-551B-1
9	1075.2	43.8	9709	4	US-09-309-572-15
10	1075.2	43.8	12494	4	US-08-935-312-13
11	1075.2	43.8	12494	4	US-08-848-760B-33
12	1075.2	43.8	15581	3	US-08-646-538-35
13	1075.2	43.8	15581	4	US-09-503-222-35
14	1065.6	43.4	9737	2	US-08-944-449-7
15	1065.6	43.4	9737	4	US-09-353-362-7
16	1030.4	41.9	9746	1	US-08-022-835-3
17	1030.4	41.9	9746	1	US-08-388-809-3
18	1030.4	41.9	9746	2	US-08-647-714-3
19	1029.2	41.9	8932	4	US-09-124-900-1
20	1029.2	41.9	8933	3	US-08-463-210-4
21	1029.2	41.9	8933	4	US-09-620-958A-3
22	1029.2	41.9	8933	4	US-09-620-958A-9
23	1029.2	41.9	8933	4	US-09-620-958A-9
24	1016.8	41.4	9739	1	US-08-022-835-1
25	1016.8	41.4	9739	1	US-08-388-809-1
26	1016.8	41.4	9739	1	US-08-647-714-1
27	1011.6	41.2	5362	3	US-08-463-210-5

28	999.8	40.7	3033	2	US-07-743-357-19	Sequence 19, Appl
29	999.8	40.7	4113	2	US-07-743-357-21	Sequence 21, Appl
30	994	40.5	9207	3	US-08-388-353-800	Sequence 800, App
31	994	40.5	9207	3	US-08-488-551B-800	Sequence 800, App
32	904.2	36.8	3856	2	US-07-743-357-20	Sequence 20, Appl
33	898.8	36.6	2739	2	US-07-743-357-16	Sequence 16, Appl
34	818.4	33.3	2348	4	US-08-876-546A-1	Sequence 1, Appl
35	818.4	33.3	2348	4	US-08-412-252-1	Sequence 1, Appl
36	818.4	33.3	2348	4	US-09-079-675-1	Sequence 1, Appl
37	805.8	32.8	9793	1	US-08-470-202-56	Sequence 56, Appl
38	805.8	32.8	9793	1	US-08-471-770-56	Sequence 56, Appl
39	805.8	32.8	9793	2	US-08-468-059-56	Sequence 56, Appl
40	805.8	32.8	9793	4	US-09-109-916-56	Sequence 56, Appl
41	798.2	32.5	1680	5	PCT-US93-06748-1	Sequence 1, Appl
42	578.8	21.5	1256	4	US-09-158-695-18	Sequence 18, Appl
43	527.8	21.5	3168	2	US-08-659-251-14	Sequence 14, Appl
44	527.8	21.5	3168	4	US-09-256-490-14	Sequence 14, Appl
45	527.8	21.5	3168	5	PCT-US96-11445-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-117-217-7
; Sequence 7, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-117-217-7

Query Match	44.0%	Score 1081.6;	DB 4;	Length 2601;
Best Local Similarity	67.9%	Pred. No. 4e-169;		
Matches 161;	Conservative	0;	Mismatches 734;	Indels 28; Gaps 6;
QY	14	TGCGCGGGCCATGAGCGAGCCACCA---CGCCCAACATCCTGATGCGCGCAGCAACT	70	
Db	77	TGCGTGAAGCAATGAGCGAGTAAACAATTCAGTACCATAATGATGCGAGAGGCAATT	136	
QY	71	TCAGGGCCCCCAAGCGATCATCAAGTCTTCAACTGGCGCAAGAGGGCCACATCGGCC	130	
Db	137	TTAGGAACCAAGAAAGATTGTTAAGTCTTTCATTTGTCGCAAGAGGGCCACAGCCA	196	
QY	131	GCAACTGCGCGCCCCCGCGCAAGAGGCTCTGGAAGTGGCGCAAGAGGGCCACCAGA	190	
Db	197	GAATTCAGGGCCCCCTAGGAAAGAGGCTTTGGAATTTGGAAAGAGGAGGACCAAAA	256	
QY	191	TGAAGGACTGCACCGAGCGCGCCCAACTTCTTCGCGAGGAGCTGCGCTTCCCGCCAGG	250	
Db	257	TGAAAGATTGTACTGAGAGACAGCGCTTA-TTTTATGGGAGATCTGCGCTTCTCTACAG	315	
QY	251	GCAAGGCGCGGAGTTTCCCGAGCGAGAGAACCGCGCAACAGCCCCCACCAGCGCGAGC	310	
Db	316	GGAAGCGCGGGAATTTTCTTCAGAGCAGACGACGACCAACAGCCCCCACCAGAGAGC	375	

QY 311 TGCAGGTGCGCG- - - - -CGACAACCCCGCAGCGAGCGCGCGCGCGCGCGCGCA 364
Db 376 TTCAGGTCTGGGTAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCGGATAGACAAGAA 435
QY 365 - - - - -CCCTGAATTCGCCCGCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
Db 436 CTGTATCTTTAACTTCCTTCAGGTCACTCTTTGGCAACGACCCCTCGTCACAAATAAGA 495
QY 419 TGGCGGCCAGATCAAGAGGCGCTGTGGACACCGCGCGCGCGCGCGCGCGCGCGG 478
Db 496 TAGGGGGCAACTAAGGAAGCTCTATTAGATACAGGACGATGATACGATTAATAGAAG 555
QY 479 AGATGAGCTTGC CGGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGGCTTCATCA 538
Db 556 AATGAGTTTGGCCAGGAATGAACCAAAATGATAGGGGAATTTGAGGTTTTATCA 615
QY 539 AGTGCGCCAGTACGACAGATCCTGATCGAGATCTGCGCGCAAGAAGGCCATCGGCACCG 598
Db 616 AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGGACATAAAGCTATAGGTACAG 675
QY 599 TGCTGATCGGCCACCCCGTCAACATCATCGCGCGCAACATGCTGACCCAGCTGGCT 658
Db 676 TATTAGTAGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCTGAACTTCCCATCAGCCCATCGAGACCGCTGCCCGTGAAGCTGAAGCCCGGCA 718
Db 736 GCACCTTAAATTTCCCATTTAGCCCTATTGAGACTGTACCAGTAAATTTAAAGCCAGAA 795
QY 719 TGACGCGCCCAAGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTCAGCG 778
Db 796 TGGATGCCCCAAGTTAAACAATGGCCATTGTACAGAAGAAAATAAAGACATTAGTAG 855
QY 779 CCATCTCGAGGAGTGGAGAAGGAGGCAAGATCAACAAGATCGGCCCGCGAGAACCCCT 838
Db 856 AATTTGTACAGAGATGGAAGAGGACGGAATTTCAAAAATTTGGCCCTGGAATCCAT 915
QY 839 ACAACACCCCGTGTGGCCATCAAGAGAAGGACAGCAACCAAGTGGCGCAAGCTGGTG 898
Db 916 ACAATCTCCAGTATTTGCCATAAAGAAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 975
QY 899 ACTTCGCGAGCTGACACGCGACCCAGGACTTCTGGGAGTGCAGCTGGGATCCCC 958
Db 976 ATTTACAGAGACTTAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC 1035
QY 959 ACCCGCGCGCTGAAGAAGAAGAGCGTGCAGCGTGGAGCTGGCGGAGCGCTACT 1018
Db 1036 ATCCGCGAGCTTAAGAAAGAAAATCAGTAACAGTACTGGATGGGTGATGCATATT 1095
QY 1019 TCAGCTGCGCCCTGGAGGAGCTTCGCAAGTACACCGCTTCACATCCCGCAGCATCA 1078
Db 1096 TTTCAAGTTCCCTTAGATGAAGACTTCAGGAAGTATACCTGCAATTTACCATTAGTATAA 1155
QY 1079 ACAACGAGACCCCGCATCCGCTACCAGTACAAGCTGTGCGCCCGCGGCTGGAAGGCA 1138
Db 1156 ACAATGAGACACCGAGGATTAGATATCAGTACAAATGTGCTTCACAGGGATGGAAGGAT 1215
QY 1139 GCGCCAGCATTTCCAGAGCAGATGACCAAGATCTCGGAGGCTTCGCGCGCGCAACC 1198
Db 1216 CACCAGCAATATTTCCAAAGTAGCATGACAAAATCTTAGAGCCCTTTAGAAAACAAATC 1275
QY 1199 CCGAGATCGTGAATACCA- - - - -GGCCCCCTGTACGTGGGACGACCTGGAGATCG 1252
Db 1276 CAGACATAGTTATCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAG 1335
QY 1253 GCAGACACCCCGCAGATCGAGGCTGCGCAAGCACTGCTGCGCTGGGGCTTACCA 1312
Db 1336 GGCAGCATAGAACAAAATAGAGGCTGAGCAACATCTGTTGAGGTGGGACTTACCA 1395
QY 1313 CCCCCGACAGAACCCAGGAGCGCCCTTCTGCGCCAT- - - - -CGAGCTGCACC 1366
Db 1396 CACCAGACAAAACATCAGAAAGACCTCCATTCCTTTGGATGGGTTATGAATCCATC 1455
QY 1367 CCGACAAAGTGGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCGTGAACG 1426

Db 1456 CTGATAAATGGACAGTACAGCCTATAGTGTCCAGAAAAAGACAGCTGACTGTCAATG 1515
QY 1427 ACATCCAGAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGAGTCTACCCCGGATCAAG 1486
Db 1516 ACATACAGAAAGTTAGTGGGAAATTTGAATTTGGGCAAGTCAGATTTACCCAGGATTAAG 1575
QY 1487 TGGCCAGCTGTCAAGCTGCTGCGGGCGCCAAAGGCTTCGACCGCATCGTGGCCCTGA 1546
Db 1576 TAAGGCAATTATGTAACCTCTTAGAGGAACCAAGACCTAACAGAGTAATACCATAA 1635
QY 1547 CCAGAGAGCGCGAGCTGGAGCTGGCCGAGAACCCGAGATCTCTGCGGAGCCCGTGACG 1606
Db 1636 CAGAAGAAGCAGAGCTAGAAGTGGCAGAAAAACAGAGAGATTTCTAAAAGAAACCACTACATG 1695
QY 1607 GCCTGTACTACGACCCCGCAGCAAGGACCTGTGTGCGCGAGATCCAGAGCAGGCGCACGACC 1666
Db 1696 GAGTGTATTATGATCCCATCAAAAGACTTAATAGCAAAATACAGAGCAGGGGCAAGGCC 1755
QY 1667 AGTGGACCTACCAGATCTACAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 1756 AATGGACATATCAATTTATCAGAGCCATTTAAAATCTGAAAACAGGAATATGCCAA 1815
QY 1727 AGATGGCACCCCGCCACACACGACGCTGAAGCAGCTGACCGAGCGCTGCAGAGATCG 1786
Db 1816 GAATGAGGGGTGCCACACTAATGATGTAAGAACAAATTAACAGAGGCAAGTGCAAAATAA 1875
QY 1787 CCATGGAGACATCTGATCTGGGCAAGACCCCAAGTTCCGCTCCCATCCAGAGG 1846
Db 1876 CCACAGAAAGCATAGTAATATGGGAAAGACTCTCTAAATTTAACTCCCATCAAAAGG 1935
QY 1847 AGACTGGGAGACTGGTGGACCGACTACTTGGCAGGCGACCTGGATCCCGAGTGGGAGT 1906
Db 1936 AACAATGGGAACATGCTGGACAGAGTATTGGCAAGCCACTGGATTTCTGAGTGGGAGT 1995
QY 1907 TCGTGAACACCCCGCTCTGGTGAAGCTGTGTACAGCTGGAGAGAGAGGCCATCATCG 1966
Db 1996 TTGTTAATACCCCTTCTAGTAAATTTAGTACCAGTTAGAGAAAGAACCCCATAGTAG 2055
QY 1967 GCGCCGAGACTTCTAGCTGGCGCGCCCAACCCGAGACCAAGATCGGAAGGCG 2026
Db 2056 GAGCAGAAACCTTCTATGTAGATGGGCAGCTACAGGGAGACTAAATTTAGGAAGAGCAG 2115
QY 2027 GCTACGTGACCGACCGCGCGCGCAGAGATCGTGTAGCCCTGACCGAGACCAACCAAGCAG 2086
Db 2116 GATATGTTACTAATAGAGGAAGACAAAAGTTGTACCCCTTAACCTGACACAAACAATCAGA 2175
QY 2087 AGACCGAGCTGAGGCCATCCAGCTGGCCCTGCAGGACAGGGCAGGCGAGGTGAACATCG 2146
Db 2176 AGACTGAGTTACAGCAATTTATCTAGCTTTGCAAGGATTCGGGATTTAGAAGTAACATAG 2235
QY 2147 TGACCGACACCCAGTACGCCCTGGGCATCATCCAGGCGCCGAGCCGACAAAGCGAGAGCG 2206
Db 2236 TAACAGACTCACAATATGCAATAGGAATCATTCAGCACAACCCAGATCAAAAGTAATCAG 2295
QY 2207 AGCTGTGAACACAGATCATCGAGCAGCTGTATCAAGAGAGGAGAGGTGTACTGTAGCTGGG 2266
Db 2296 AGTTAGTCAATCAATAATAGAGCAGTTAATAAAAAGGAAGAGTCTATCTGSCATGGG 2355
QY 2267 TGCCCGCCACAGAGGCGATCGGGGCAACGAGCAGATTCGACAGCTGCTGGAGCAAGGCA 2336
Db 2356 TACGACACACAAGGAATTTGGAGGAATGAACAAGTAGATAAATTTAGTCAGTCTCGAA 2415
QY 2327 TCCGCAAGGTGCTTCTCCCTGAGCGCATCGAT 2358
Db 2416 TCAGGAAGTACTATTTTATAGTGAATAGAT 2447

RESULT 2

US-09-117-217-9

; Sequence 9, Application US/09117217

; Patent No. 6221578

; GENERAL INFORMATION:

```
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)..(489)
; OTHER INFORMATION: gag P6 (52 AA)
US-09-117-217-9

Query Match          44.0%; Score 1081.6; DB 4; Length 2601;
Best Local Similarity 67.9%; Pred. No. 4e-169;
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;

QY 14 TGGCCGAGGCATGAGCAGGCGCACCA---GCGCCACATCTCTGATCGAGCGCAACT 70
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 77 TGGCTGAAGCAATGAGCCAGTAACAATTCAGCTACCAATAATGATCGAGAGGCAATT 136
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTTCAACTCGGCAAGAGGCGCACATCGCCC 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 137 TTAGGAACCAAGAAAGATGTTAAGTGTTCATTTGTGCAAAAGAGGCGCACAGCCA 196
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 131 GCRACTGCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGAGGGCCACAGA 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 197 GAAATTCAGGGGCGCCCTAGGAAAAGGGCTGTGGAAATGTGGAAGGAAGGACACAAA 256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 191 TGAAGGACTGCGACGAGCGCGCCCAACTTCTTCGCGGAGGACCTCGCCCTTCCCCAGG 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 257 TGAAGATTTGACTGAGACAGAGCTAA-TTTTTTGAAGGAGATCTGGCCCTTCTACAG 315
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 251 GCRAAGCGCGGAGTTCCCGAGCGAGCAGACACCGCGCCACAGCCCAACAGCGCGAGC 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 316 GGAAGGCGAGGGAATTTTCTCAGAGCAGACAGAGCAACAGCGCCCAACAGAGAGAGC 375
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCAGGCA 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 376 TTCAGGTCTGGGTGAGAGACAACACTCCCCCTCAGAAGCAGGAGCGGATAGACAAGGA 435
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 365 -----CCCTGAACCTCCCGCAGATCACCTGTGGCAGCGCCCGCTGTGAGCATCAAGG 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 436 CTGTATCTTTAACTTCCCTCAGTCACTCTTTGGCAACGACCCCTCGTCACATAAGA 495
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 419 TGGCGGCGCAGATCAAGGAGGCGCTGTGGACACCGCGCGCGAGCAGCACCGTGTGGAGG 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 496 TAGGGGGCAACTAAGAGAGCTCTATTAGATACAGAGCAGATGATACAGATTAGAAG 555
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 479 AGATGAGCCCTGCGCGCAAGTGGAAAGCCCAAGATGATCGCGGATCGCGGCTTCATCA 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 556 AATGAGTTTGGCAGGAAGATGGAACCCAAAATGATAGGGGGAATTTGAGGTTTATCA 615
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 539 AGGTGCGCCAGTAGCAGCATCTGTATCGAGATCTCGGCGCAAGAGGCCATCGGCACCG 598
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 616 AAGTAACACAGATGATCAGATCACTCATAGNAATCTGTGCACATAAGACTATAGTACAG 675
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 599 TGTGATGCGGCCCAACCGGTGAACATCATCGCGCCGCAACATGCTGACCCAGCTGGGCT 658
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 676 TATTAGTAGGACCTTACACCTGTCAACATAATTGGAAGAAATCTCTTGACTCAGATTGGTT 735
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 659 GCACCTGAACCTTCCCATCAGCCCCATCAGACCGGTGCGCGTGAAGCTGAAGCCCGCA 718
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 736 GCACCTTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAAGTAAATTAAGCCAGGA 795
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 719 TGGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGGCCCTGACC 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 796 TGGATGCGCCCAAGATTAAACAATGGCCATTGACAGAAGAAAATAAAGCATTAGTAG 855
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 779 CCATCTGGAGAGATGAGAAAGGAGGCAAGATCACCAAGATCGGCCCGGAGACCCCT 838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 856 AAATTTGTACAGAGATGGAAGGAGGAAATTTCAAAAATTTGGGCTGAAAATCCAT 915
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 839 ACAACACCCCGCTTGGCCATCAAGAGAAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 916 ACAATCTCCAGTATTTGCCATAAGAAAAAGACAGTACTAAAATGGAGAAAATTAGTAG 975
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 899 ACTTCCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGAGGTGCAGCTGGGCATCCCCC 958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 976 ATTTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTCAATTAGGAATACCAC 1035
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 959 ACCCGCGCGGCTGAAGAAGAAGAGCGTGAACGCTGTGAGCGTGGGCGACGCTACT 1018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1036 ATCCCGCAGGCTTAAAAAAGAAAAATCAGTAACAGTACTGGATGTGGTGTGATGATATT 1095
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1019 TCAGCGTGCCTCGACGAGGACTTCCCAAGTACACCGCCTTCACCATCCCGACATCA 1078
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1096 TTTGAGTTTCCCTTAGATGAAGACTTCAGGAAGTATACGATTTACCATACCTAGTATAA 1155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1079 ACAACGAGACCCCGGATCGCTACCAAGTACAAAGTGTCTGCCCGAGGCTGGAAGGCA 1138
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1156 ACAATGAGACACAGGATTAGATATCAGTACATGTCTTCCACAGGATGGAAGGAT 1215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1139 GCGCCAGATCTTCCAGAGCAGCATGACCAAGATCTTGGAGCCCTTCCCGCCCGCAACC 1198
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1216 CACCAAGCAATTTCCAAAGTAGCATGACAAAAATCTTTAGAGCCTTTTGAAGAAACAAATC 1275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1199 CCGAGATCGTATCTACCA-----GGCCCCCTGTAGTGGGCGAGCACCTGGAGATCG 1252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1276 CAGACATAGTTATCTATCAATACATGATGATTTGTGTAGGATCTGAGTTAGAAATAG 1335
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1253 GCCAGCAGCGCGCCAAAGATCGAGGAGCTGCGCAAGCAGCTGTGCGGTGGGCTTCACCA 1312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1336 GGCAGCATAGAACAAAAATAGAGGAGCTGACAAACATCTGTTGAGTGGGACTTACCA 1395
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1313 CCGCGCACAGAAGACCAAGAGGAGCGCCCTTCTGCTGCCAT-----CGAGTGCACC 1366
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1396 CACCAAGCAAAAAACATCAGAAAGAACCTCCATCTCTTTGGATGGTTTATGAATCCCATC 1455
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1367 CCGACAAGTGACCGTGCAGCCCATCGAGCTGCCGAGAGAGAGCTGGACCTGAACG 1426
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1456 CTGATAATGACAGTACAGCCTATAGTGTGCCAGAAAAAGACAGCTGGACTGTCAATG 1515
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1427 ACATCCAGAAGCTGGTGGCAAGCTGAACCTGGGCGAGCCAGATCTACCCCGCATCAAGG 1486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1516 ACATACAGAAGTTAGTGGGAAATTAATTGGCAAGTCAAGTTTACCAGGGATTAAG 1575
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1487 TGCCCGAGCTGTCAAGCTGTGCGCGCGCCAGAGCCCTGACCGACATCTGTGCCCTTGA 1546
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1576 TAAGCAATTTATGTAATCTCTAGAGGAACCAAGCACTAACAGAAATAATACCACTAA 1635
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1547 CCGAGAGGCGAGCTGGAGCTGCGCGAGACCGGAGATCCTGCGCAGCCCTGTCAGC 1606
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1636 CAGAAGAGCAGCTAGACTGGCAGAAAAACAGAGATTTCTAAAAGAACCAAGTACATG 1695
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1607 GCGTGTACTAGCAGCCCGAGGACCTGTTGGCGGAGATCCAGAAAGCAGGCGCACAGC 1666
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1696 GAGTGTATTATGACCCATCAAAAGACTTAATAGCAGAAATACAGAGCAGGGGCAAGGCC 1755
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1667 AGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1756 AATGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAACAGGAAATATGCAA 1815
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1727 AGATGCGCACCGCCACCAACGACGTGAAGCAGCTGACCGAGCCGTGAGAGATCG 1786
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1816 GAATGAGGGGTGCCACACTAATGTATTAACCAATTAACAGAGGCGATGCAAAAAATAA 1875
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1787 CCATGGAGATCGTGTATCTGGGCAAGACCCCAAGTTTCCGCTGCCATCCAGAAAG 1846
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

1876 CCACGAAAGCATAGTAATAATGGGAAAGACTCTAAATTTAAACTGCCCCATACAAAGG 1935
1847 AGACCTGGGAGACCTGGTGACCGACTACTGCGAGGCCACCTGGATCCCGCAGTGGGAGT 1906
1936 AACAATGGGAAACATGGTGACAGATATTGGCAAGCCACCTGGATTCCTGAGTGGGAGT 1995
1907 TCCTGAACACCCCGCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCCATCATCG 1966
1996 TTGTTAATACCCCTCCCTTAGTGAATATTGGTACCAGTTAGAGAAAGACCCATAGTAG 2055
1967 GCGCCGAGACCTTCTACGTGGAGCGGCGCCCAACCGCGAGACCAAGATCGCAAGGCCG 2026
2056 GAGCAGAAACCTTCTATGTAGATGGGCGACCTAACAGGGAGACTAAATTAGGAAAGSCAG 2115
2027 GCTACGTGACCGAGCGGGCGCGCAGAGATCGTGAAGCTGCGAGCCGTGCGAGACCAACACAGA 2086
2116 GATATGTTACTAATAGAGGAAGCAAAAGTTGTCAACCTTAAGTACCAACAAATCAGA 2175
2087 AGACCGAGCTGCGAGGCATCCAGCTGGCCCTGCAGGACACGCGGAGGTGAACATCG 2146
2176 AGACTGAGTTACAGCAATTTATCTAGCTTTGCAGGATTCGGGATTTAGAGTTAAACATAG 2235
2147 TGACCGACACAGTAGTACGCTGGCCCTGGGCATCATCCAGGCCCGCAGCAAGAGCGAGCG 2206
2236 TAACAGACTCACAAATATGCATTAGGAATCATTTCAAGCACAACCCAGATCAAAGTGAATCAG 2295
2207 AGCTGTGAACACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2266
2296 AGTTAGTCAATCAATTAATAGACAGTAAATAAAGGAAAGGTTCTATCTGGCATGGG 2355
2267 TGCCCGCCACAAAGGGCATCGCGCGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGGCA 2326
2356 TACCAGCACAAAGGAATTTGGAGGAATGCAACAAGTAGATAATTAGTCAGTGTCTGGAA 2415
2327 TCCGCAAGTGCTGCTTCCTGGAGCGGATCGAT 2358
2416 TCAGAAAGTACTATTTTATAGATGGAATAGAT 2447
```

RESULT 3

```

US-09-117-217-11
; Sequence 11, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kust
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (453)...(749)
; OTHER INFORMATION: Protease
US-09-117-217-11
```

```

Query Match 44.0%; Score 1081.6; DB 4; Length 2601;
Best Local Similarity 67.9%; Pred. No. 4e-169;
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;

QY 14 TGCCCGAGCCATGAGCCAGGCCACCA---GCGCCACATCTGTGTCAGCGCAGCGCAACT 70
DB 77 TGGCTGAAGCAATGAGCCCAAGTAACAAATTCAGCTACCAATAATGATGCAGAGGCAAT 136
```

```

QY 71 TCAGGSCCCCAAGCGCATCATCAAGTGTCTCAACTGCGCAAGGAGGCCACATCGCCC 130
DB 137 TTAGGAACCAAAAGAAAGATTGTTAAGTGTTCATTTGTCGCAAAAGAGGCCACACGCCA 196
QY 131 GCAACTGCGCCGCGCCCGCGCAAGAGGGCTGCTGGAAGTGCGGCAAGGAGGCCACACAGA 190
DB 197 GAAATTCGACGGCCCTTAGCAAAAAGGGCTGTGGAATGTGGAAGAGGAAGCACACAAA 256
QY 191 TGAAGGACTCGACCGAGCGCCAGCGCAACTTCTTCCCGGAGGACCTGGCTTCCCCCAGG 250
DB 257 TGAAGATTTACTGAGAGACAGGCTAA-TTTTATTAGGAAGATCTCGCCCTTCCTAGCAAG 315
QY 251 GCAAGGCGCCGAGTTCGCCAGCGAGCAGAACCGCGCCACAGCCGCCACCGCGCGCAGC 310
DB 316 GGAAGGCGAGGAAATTTCTTCAGAGCAGACAGAGCCACAGCCGCCACCAAGAGAGAG 375
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCAGCGCGCAGGGCA 364
DB 376 TTCAGGTCTGGGTAGAGACAACTCCCTCCCTCAGAAGCAGGAGCGGATAGACAAGAA 435
QY 365 -----CCCTGAACTTCCCGCAGATCACTCTGTGGCAGCGCCCTGTGTGAGCATCAAG 418
DB 436 CTGTATCTCTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTGTCACATAAAGA 495
QY 419 TGGCGCGCCAGATCAAGGAGGCCCTGCTGCACACCGCGCGCAGCACACCGTGTCTGCAGG 478
DB 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555
QY 479 AGATGAGCCTTCCCGCGCAAGTGGAAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 556 AAATGAGTTTGCAGGAAGATGGAACCAAAATGATAGGGGAATTTGGAGGTTTTATCA 615
QY 539 AGTGCGCAGTAGGACCATCTCTGATCAGATCTCGCGCAAGAGGCCATCGGCACCG 598
DB 616 AAGTAAGACAGATGATGATCAGATCTCATAGAAATCTGTGGACATATAAAGCTATAGGTACAG 675
QY 599 TGCTGATCGCGCCCGCACCCCGTGAACATCATCGGCCCAACATGCTGACCCAGCTGGGCT 658
DB 676 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACTCAGATTGGTT 735
QY 659 GCACCTGAACTTCCCGCATCAGCCCCCATCGAGACCGGTGCGCGTGAAGCTGAAGCCCGCA 718
DB 736 GCACCTTAAATTTTCCCATTAGCCCTATTGAGACTGTACCAGTAAATTTAAAGCCAGAA 795
QY 719 TGAAGCGCCCAAGGTGAACAGTGGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 796 TGGATGCGCCCAAGGTAAACAATGGCCATTGACAGAAGAAAAATAAAGCATTTAGTAG 855
QY 779 CCATCTGCGAGGAGATGGCAAGGAGGCGCAAGATCACCAAGATCGGCCCGCGAGAACCCCT 838
DB 856 AATTTGTACAGAGATGGAAGAGGAGGAGAAATTTCAAAATTTGGCCCTTGAATCCAT 915
QY 839 ACAACACCCCGCTGTTGCGCATCAAGAAGAGGACAGCACCAAGTGGCGCAAGCTGGTGG 898
DB 916 ACAATACTCCAGTATTGGCCATAAAGAAAAAGACAGTACTAAATGGAGAAAAATTAGTAG 975
QY 899 ACTTCCGCGAGCTGAACAACCGCCAGGACTTCTGGGAGGTGCGAGCTGGGATCCCCC 958
DB 976 ATTTTCAGAGAACTTAATAGAGAACTCAAGACTTCTGGGAAGTTCAATTTAGGAATACAC 1035
QY 959 ACCCGCGCGCCCTGAAGAAGAAAGAGCGCTGACCGTGTGCGAGCTGGGCGGACCGCTACT 1018
DB 1036 ATCCCGCAGGCTTAAAAAGAAAAATCAGTAACAGTACTGGATGTGGGTGATGCATATT 1095
QY 1019 TCAGCGTGCCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCCACATCCCCAGCATCA 1078
DB 1096 TTTCACTTCCCTTTAGATGAAGACTTCAGGAAGTATCTGCACTTACCATACCTAGTATAA 1155
QY 1079 ACAACGAGACCCCGCATCCGCTACAGTACACGCTGCTGCCCCCGGCTGGAGGGCA 1138
DB 1156 ACAATGAGACACCGAGGATTAGATATCAGTACAAATGTGCTTCCACAGGGATGGAAGGAT 1215
```

QY 1139 GCCCGACGATCTTCCAGACAGCATGACCAAGATCTCGAGCCCTTCGGCGCCGCAAC 1198
Db 1216 CACCAGCAATATTCCAAAGTAGCATGACAAAATCTTAGAGCCTTTTAGAAAACAAATC 1275
QY 1199 CCAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGAGGACCTGGAGATCG 1252
Db 1276 CAGACATAGTTATCTATCAATACATGATGATTTGTATGTAGATCTGACTTAGAATAG 1335
QY 1253 GCCAGCACCGCGCAGATCGAGGAGTGGCGAAGCAGCCTGCTGCGTGGGGTTCACCA 1312
Db 1336 GGAGCATAGAACAAAATAGAGAGCTGAGACAACATCTGTTGAGGTGGGACTTACCA 1395
QY 1313 CCCCCGACAGAAACACCAAGAGGAGCCGCCCTTCCTGCCCAT-----CGAGCTGCAC 1366
Db 1396 CACCAGACAAAACATCAGAAAGAACCTCCATTCTTTGGATGGTATGAACCTCCATC 1455
QY 1367 CCAGCAAGTGGACCGTGCAGCCCATCGAGTGGCCGAGAGGAGAGCTGACCGTGAACG 1426
Db 1456 CTGATAAATGGACAGTACAGCCTATAGTGTGCTGCCAGAAAAGACAGCTGGACTGTCAATG 1515
QY 1427 ACATCCAGAAAGCTGGTGGCAAGCTGAACCTGGCCAGCCAGATCTACCCGGCATCAAGG 1486
Db 1516 ACATACAGAAAGTGTAGTGGGAAATTTGAATTTGGCAAGTCAAGTTTACCCAGGATTAAG 1575
QY 1487 TGGCCAGCTGTGCAAGCTGCTGCGGGCCGCAAGGCCCTGACCGACATCGTCCCTGTA 1546
Db 1576 TAAGGCAATTATCTAACTCTTAGAGGAACCAAGCACCTAACAGAAGTAATACCACATA 1635
QY 1547 CCAGAGGCGGAGCTGGAGCTGGCGAGAACCCGCGAGATCTCTGGCGGAGCCCGTGACG 1606
Db 1636 CAGAAGAAGCAGAGCTAGACTGGCAGAAAACAGAGAGATCTTAAAGAACCAAGTACATG 1695
QY 1607 GCCTGTACTACGACCGCAGCAAGCAAGCTGCTGGCCGAGATCCAGACGACGAGCCAGAC 1666
Db 1696 GAGTGTATTATGACCATCAAAAGACTTAATAGCAAAATACAGAGAGCGGGGCAAGGCC 1755
QY 1667 AGTGGACCTACCAAGATCTACCGAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
Db 1756 AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAACAGGAAATATGCAA 1815
QY 1727 AGATGGCACCGCCGACCAACAGAGCTGGAAGCAGCTGACCGAGGCGGTGCAGAAATCG 1786
Db 1816 GAATGAGGGTGGCCCACTAAATGATGTAAACAATTAACAGAGGAGTGCAGAAAATAA 1875
QY 1787 CCATGGAGACATCGTGATCTGGGCAAGACCCCAAGTTCGCCCTGCCATCCAGAAGG 1846
Db 1876 CCACAAAGCATAGTAATATGGGGAAGACTCTTAATTTAACTGCCCCATCAAAAGG 1935
QY 1847 AGACCTGGGAGACCTGGTGACCGACTACTTGGCAGGCCAGCTGGATCCCGAGTGGGAGT 1906
Db 1936 AAACATGGGAAACATGGTGACAGAGATATTGGCAAGCCACTGGAATTCCTGAGTGGAGT 1995
QY 1907 TCGTGACACCCCGCCCTGGTGAAGTGTGGTACAGCTGGAGAGGAGGCCCATCATCG 1966
Db 1996 TTGTTAATACCCCTCCCTTAGTGAATATATGTTACCACTTACCAAGTACCAAGG 2055
QY 1967 GCGCCGAGACCTTCTAGTGGCGCGCCGCAACCCGAGACCAAGATCGCAAGGCCG 2026
Db 2056 GACAGAAACCTTCTATGTAGTGGGAGCTAACAGGGAGCTTAATAGGAAGAGCAG 2115
QY 2027 GCTACGTGACCGACCGCGGCGCAGAAAGATCGTGAGCTTGAACCGAGACCAACCAACAGA 2086
Db 2116 GATATGTACTAATAGAGGAAGACAAAAGTTCTCACCCTTAACAGACAAACAATCAGA 2175
QY 2087 AGACCGAGCTGCAGGCATCCAGCTGGCCCTGCGAGGACGCGGAGCGAGGTGAACATCG 2146
Db 2176 AGACTGAGTTTACAAAGCAATTTATCTAGCTTTGAGGATTCGGGATTAGAAGTAAACATAG 2235
QY 2147 TGACCGACAGCCAGTACGCCCTGGGATCATCCAGGCCAGCCGACAGAGCGAGAGCG 2206
Db 2236 TAACAGACTCACAAATATGCAATTAGGAATCATTCAGCACCAACAGATCAAAGTGAATCAG 2295
QY 2207 AGCTGTGAACCAAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGG 2266

Db 2296 AGTTAGTCAATCAATAATAGACAGTTAATAAAAGGAAAGGTCTATCTGGCATGGG 2355
QY 2267 TGCCCCCCCACAAGGCGATCGGGCAACGACGAGATCGACAAGCTGGTGCAGCAGGGCA 2326
Db 2356 TACCAGCACACAAGGAATTTGGAGGAAATGAACAAGTAGATAAAATTAGTCAGTCTGGAA 2415
QY 2327 TCCGCAAGGTGCTGTTCTCTCGACGCGCATCGAT 2358
Db 2416 TCAGGAAGTACTATTTTAGATGGAATAGAT 2447

RESULT 4
US-09-117-217-13
; Sequence 13, Application US/09117217
; Patent No. 6221578
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; TITLE OF INVENTION: OF HUMAN HIV STRAINS
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/117,217
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (750)..(2435)
; OTHER INFORMATION: Reverse Transcriptase
US-09-117-217-13

Query Match 44.0%; Score 1081.6; DB 4; Length 2601;
Best Local Similarity 67.9%; Pred. No. 4e-169;
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;
QY 14 TGGCCGAGCGCATGAGCAGCGCCACCA---GCGCCAAACATCTCTGATCGAGCGCAGCAACT 70
Db 77 TGGCTGAAGCAATGAGCCAGTACAAATTCAGCTACCAATATGTCAGAGAGGCAATT 136
QY 71 TCAGGCGCCCAAGCGCATCATCAAGTCTTCAACTCGGCAAGGAGGCGCACATCGCC 130
Db 137 TTAGGAACCAAGAAAGATTGTTAAGTGTTCATTTGTCGCAAGAGGGGCACACAGCA 196
QY 131 GCAACTCCCGCGCCCGCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGGCCACCAGA 190
Db 197 GAAATGTCAGGGCGCCCTAGGAAAAGGGCTGTGGAAATGTGGAAGAGGAGGACACCAA 256
QY 191 TGAAGGACTCCACGCGCGCCGCAAGGCAACTTCTTCCGAGGACCTGGCCCTTCCCGCAGG 250
Db 257 TGAAGATTGCTACTGAGAGACAGGCTAA-TTTTATGGAGAGATCTGGCCCTTCTTACAAG 315
QY 251 GCAAGGCGCGGAGTTCGCCAGGAGAGCAACCGCCCAACAGCCCAACAGCCCGCCGAGC 310
Db 316 GGAAGGCGAGGGAATTTCTTCAGAGCAGACACAGAGCCACAGCCCGCCAGAGAGAGC 375
QY 311 TGCAGGTGCGCGG-----CGACAAACCCCGCAGCGAGGGCGGGCGCGGCGCCAGGGCA 364
Db 376 TTCAGGTCTGGGTGAGAGACAACAACCTCCCGCTCAGAAAGCAGGAGCCGATAGACAAGGA 435
QY 365 -----CCCTGAATCTCCCGCAGATCACCTGTGGCAGCGCCCTCTGTGAGCATCAAG 418
Db 436 CTGTATCTTTAACTTCCCTCAGGTCACTCTTTGGCAACGACCCCTCTGTACATAAGA 495
QY 419 TGGCGCGCCAGATCAAGGAGGCGCTTGTGGACACCGCGCGGCGAGCACCGTGTGGAGG 478
Db 496 TAGGGGGCAACTAAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG 555

; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-388-353-1

Query Match 43.8%; Score 1075.2; DB 3; Length 9709;
Best Local Similarity 67.7%; Pred. No. 5.1e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGCCGAGGCCATGAGCCAGGCCACCA---GCGCCACATCCTGTGTCAGCGCAGCAACT 70
DB 1877 TGGCTGAAGCAATGAGCCAAAGTAACAAATCCAGCTACCATTAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGGCCACATCGCCC 130
DB 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATATTGTGCAAGAGGCGCATAGCCA 1996
QY 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGGCAAGAGGGCCACACAGA 190
DB 1997 AAAATTGTCAGGGCCCTAGGAAAAGGGCTGTTGGAATGTGAAAAGGAAGGACACCAA 2056
QY 191 TGAAGACTCACCGAGCGCCAGGCCAACTTCTCCCGAGGAGCTGGCCCTTCCCCCAGG 250
DB 2057 TGAAGATTGCTGAGAGACAGGCTAA-TTTTTATGGGAAGATCTGGCCCTTCCCACAG 2115
QY 251 GCAAGGCCCGCGATTCTCCCGAGCGAGCAGAAACCGCGCAACAGCCCAACAGCGCGGAGC 310
DB 2116 GGAAGGCCAGGGAATTCTTCAGAGCAGACAGACAGCCACAGCCCAACAGAGAGAGC 2175
QY 311 TGCAGTGCACGG-----CGAACCCCGCAGCGAGGCGCGCGCGCGCGCGCGCGCA 364
DB 2176 TTCAGGTTTGGGAAGAGACAACAACTCCCTCTCAGAAGCAGGAGCGGATFAGACAAGAA 2235
QY 365 -----CCCTGAACTTCCCGCAGATCACCTGTGGCAGCGCCCTGTGTCAGCATCAAG 418
DB 2236 CTGTATCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCAACAATAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGG 478
DB 2296 TAGGGGGCAATTAAAGGAAGCTCTATTAGATACAGAGCAGATGATACAGATTAGAAG 2355
QY 479 AGATGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
DB 2356 AAATGAATTTGCCAGGAAGATGGAACCAAAAATGATAGGGGAAATTGGAGGTTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACAGATCCTGTATCGAGATCTGCGCAAGAGAGGCCATCGCACCG 598
DB 2416 AAGTAGACAGTATGATCAGATCTCATAGAAATCTGCGGACATAAAGCTATAGGTACAG 2475
QY 599 TGTGTATGCGCCCAACCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACCTTCCCATCAGCCCATCGAGACCGTGCAGCGGATCGCGCGCGGCA 718
DB 2536 GCACCTTAAATTTTCCCATTTAGTCTATTGAGACTGTACCAAGTAAATTTAAAGCCAGAA 2595
QY 719 TGGACGGCCCAAGGTGAAGCTGCGCCCTGACCGAGGAGAGATCAAGGCCCTGACCG 778
DB 2596 TGGATGCCCCAAAGTTAAACAATTGGCCATTGACAGAAGAAAATAAAGCAATTAGTAG 2655
QY 779 CCATCTGCGAGAGATGGAGAAGGAGGCAAGATCATCAAGATCGCGCCCGGAGAACCCCT 838
DB 2656 AAATTGTACAGAAATGGAAGAGGAGGAAAATTTCAAAAATTTGGCCCTGAAAATCCAT 2715
QY 839 ACAACACCCCGTGTTCGCCATCAAGAAAGGACAGCAACCAAGTGGCGCGCAAGCTGGTG 898

DB 2716 ACAATACTCCAGTATTTCATATAAGAAAAAGACAGTACTATAATGAGAAAAATTAGTAG 2775
QY 899 ACTTTCGCGAGCTGAACAAGCCAGCCAGGACTTCTGGGAGGTGCAGCTGGGCATCCCC 958
DB 2776 ATTTAGAGAACTTAATAGAGAACTCAAGATTTCTGGGAAGTTCAATTTAGGAATACAC 2835
QY 959 ACCCCGCGCCCTGAAGAAGAAAGACGCTGACCGTGTGGAGGTGGGGGAGCCCTACT 1018
DB 2836 ATCCTGCAGGTTAAACACAGAAAAATCAGTAACAGTACTGGATGTGGCGCATGATATT 2895
QY 1019 TCAGCTGCCCCCTGGAGGAGACTTCCGCAAGTACACCGCTTCACCATCCCCAGCATCA 1078
DB 2896 TTTCAAGTTCCTTAGATAAAGACTTTCAGGAAGTATACTGCAATTTACCATACCTAGTATA 2955
QY 1079 ACACAGAGACCCCGCATCCGTACCAAGTACACAGTGTCTGCCCGAGGGCTGGAAGGCA 1138
DB 2956 ACAATGAGACACCGAGGATTAGATATCAGTACAATGTGCTTCCACAGGATGGAAGGAT 3015
QY 1139 GCGCCAGCATCTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGGCGCCGCAACC 1198
DB 3016 CACCAGCAATATTCCAGTGTAGCATGACAAAAATCTTTAGAGCCTTTTAGAAAAACAATC 3075
QY 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTGTACGTGGCGAGCGACCTGGAGATCG 1252
DB 3076 CAGACATAGTCACTATCAATACATGATGATTGTTGTATGTAGGATCTGACTTAGAAATAG 3135
QY 1253 GCAGACACCGCGCAGATCGAGGAGCTGCGAAGCACTGCTGCGCTGGGGCTTCACCA 1312
DB 3136 GGAGCATAGAACAAAAATAGAGAACTGAGACAACATCTGTTGAGTGGGATTTACCA 3195
QY 1313 CCCCCACAAAGACACCAAGAGCGCCCCCTTCTTCCCTCCCAT-----CGAGCTGCACC 1366
DB 3196 CACCAGACAAAAACATCAGAAAGAACTCCATTTCTTGTGATGGGTTATGAATCCATC 3255
QY 1367 CCGAAGAGTGGCGTGCAGCCCATCGAGCTGCCGAGAGAGAGAGCTGGACCGTGAACG 1426
DB 3256 CTGATAAATGGCAGTACAGCCTATAGTGTGCCAGAAAAAGCAGCTGGACTGTCAATG 3315
QY 1427 ACATCAGAAAGCTGGTGGCAAGCTGAACTGGCGACAGATCTACCCCGCATCAAG 1486
DB 3316 ACATACAGAAATTTAGTGGGAAATTTGAATTTGGCAAGTCAAGATTTATGACGGGATTAAG 3375
QY 1487 TGGCCAGCTGTCAAAGCTGCTGCGGGGCCCAAGGCCCTGACCGACATCGTGCCCTGA 1546
DB 3376 TAAGGCAATTTATGAACTTCTTAGGGGAACCAAGCACTAACAGAAAGTAGTACCACCTAA 3435
QY 1547 CCAGAGGCGCGAGCTGGAGCTGGCCGAGAACCCGAGAGATCTTCCGCGAGCCCGTGCAG 1606
DB 3436 CAGAAAGACGAGCTAGAACTGGCAGAAAACAGGGAGATTCTAAAAAGAACCGGTACATG 3495
QY 1607 GCGTGTACTACGACCCCGCAGCAAGCCCTGCTGGCCGAGATCCAGAACGAGGCCACGACC 1666
DB 3496 GAGTGTATTATGACCCCATCAAGAGCTTAATAGCAGAAATACAGAACGAGGGCAAGGCC 3555
QY 1667 AGTGGACCTTACCAGATCTACCAGAGCCCTTCAAGAACCTCAAGACCCGCAAGTACGCCA 1726
DB 3556 AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGGAAAAATATCAA 3615
QY 1727 AGATGCGCACCCGCCACCAACGACGTGAAGCAGCTGACCGAGGCGCGTGCAGAGATCG 1786
DB 3616 GAATGAAGGGTGGCCACACTAATGATGTGAACAATTTAACAGAGGCGAGTACAAAAATAG 3675
QY 1787 CCATGGAGACATCTGATCTGGGGCAAGACCCCAAGTTTCCGCTCCCATCCAGAAAG 1846
DB 3676 CCACAGAAACATAGTAATATATGGGAAAGACTCTTAATTTAAATTTACCATACAAAGG 3735
QY 1847 AGACCTGGGAGACCTGGTGGACCGACTACTTGGCAGCGACCTGATCCCCAGTGGAGT 1906
DB 3736 AAACATGGGAAGCATGGTGGACAGAGTATTGGCAAGCCACTGGATTTCTTCTGAGTGGAGT 3795
QY 1907 TCGTGAACACCCCGCCCTGCTGTAAGCTGTGTTACCAAGCTGGAGAGGAGGCCCATCATCG 1966

Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGTTACCAAGTTAGAGAAAGAACCCATAATAG 3855
QY 1967 GCGCGGAGACCTTCTAGCTGAGCGCGCCGCAACCGCGAGAACCAAGATCGGCAAGCCG 2026
Db 3856 GAGCAGAACTTTCTATGTAGATGGGAGCCCAATAGGGAACCTAAATTAGGAAAGCAG 3915
QY 2027 GCTACGTGACCGAGCGGGCGGCGAGAGATCGTGAGCTTACCGGAGACCAACCAACAGA 2086
Db 3916 GATATGTAACCTGACAGAGGAGACAAAAGTTGTCCTCCCTTAACCGACACACAAATCAGA 3975
QY 2087 AGACCGAGCTGACAGGCATCCAGCTGGCCCTGCGAGGACGCGGAGGAGTGAACATCG 2146
Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTTCAGGATTCGGGATTTAGAAATAACATAG 4035
QY 2147 TGACCGACAGCAGTACGCTCGGCTGCGCATCTCCAGGCGCCGAGCGGCAAGAGCGAGCG 2206
Db 4036 TGACGAGCTACATATGCTATGGAATCATCTCAAGCACACACCAAGATGAAGTGAATCAG 4095
QY 2207 AGCTGTGAACAGATCATCGAGCAGCTGATCAAGAAGGAGAAAGTGTACCTGAGCTGGG 2266
Db 4096 AGTTAGTCAGTCAATATAGAGCAGTTAATAAAGGAGGAAAGTCTACCTGCGCATGGG 4155
QY 2267 TGCCCGCCCAAGGAGCATCGGCGGCAACGAGCAGATCGACAGCTGTTGAGCAAGGCGCA 2326
Db 4156 TACAGCACACAAAGAAATTTGGAGGAAATGAACAAGTAGATGGGTTGTCAGTCTGGAA 4215
QY 2327 TCCGCAAGGTGCTTCTCTCGACGCGCATCGAT 2358
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 8
US-08-488-551B-1

; Sequence 1, Application US/08488551B
; Patent No. 6015661
; GENERAL INFORMATION:
; APPLICANT: Nicholas J. Deacon
; APPLICANT: Dale A. McPhee
; APPLICANT: David Cooper
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 841
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 11530-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,551B
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PM3864 (AU)
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: PM4002 (AU)
; FILING DATE: 21-FEB-1994
; APPLICATION NUMBER: PM0284 (AU)
; FILING DATE: 23-DEC-1994
; APPLICATION NUMBER: US 08/388,353
; FILING DATE: 14-FEB-1995
; APPLICATION NUMBER: PM3021/95
; FILING DATE: 17-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: FRANK S. DIGILIO
; REFERENCE/DOCKET NUMBER: 9606Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9709 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-488-551B-1

Query Match 43.8%; Score 1075.2; DB 3; Length 9709;
Best Local Similarity 67.7%; Pred. No. 5.1e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGGCCGAGGCCATGAGCCAGGCCACCA---GCGCCAAACATCTGTATGATGAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCCAAGTAACAAATCAGCTACCATTAATATACATAAGGAAGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTCTTCACTGCGGCAAGGAGGCCACATCGGCC 130
Db 1937 TTAGAACCAAGAAGAAAGACTGTTAAGTGTTCATTTGTCGCAAGAGGGCACATAGCCA 1996
QY 131 GCAACTGCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGAGGGCCACCAAGA 190
Db 1997 AAAATTGCGAGGCCCTAGGAAAAGGGCTGTGGAAATGTGAAATGTGAAAGGAAGGACACCAA 2056
QY 191 TGAAGGACTGACCGAGCGCCAGGCCCAACTTCTCCCGGAGGACCTGGCTTCCCGCAGG 250
Db 2057 TGAAGATTTGACTGAGAGACAGGCTAA-TTTTATAGGGAAGATCTGGCTTCCCAACAG 2115
QY 251 GCAAGGGCCGCGAGTTCCCGAGCGAGACAACCGCCCAACAGCCCAACAGCCCGCGAGC 310
Db 2116 GGAAGCCAGGGAATTTCTTCAGAGCAGACAGCCCAACAGCCCAACAGAGAGAGC 2175
QY 311 TGCAGTGGCGGG-----CGACAACCCCGAGCGCGCGCGCGCGCGCGCGCGCGCA 364
Db 2176 TTCAGTTTGGGGAAGAGACAACAACCTCCCTCTCAGAAGCAGGAGGCCCATAGACAAGAA 2235
QY 365 -----CCCTGAACCTCCCGCAGATCACTCTTGGCAGCGACCCCTCGTCAACAATA 2295
Db 2336 CTGTATCTTTAGCTTCCCTCAGATCACTCTTGGCAGCGACCCCTCGTCAACAATA 2295
QY 419 TGGCGCGCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCGCGCGCGCGCGCGCGCG 478
Db 2296 TAGGGGCAATTAAGGAAGCTCTATTAGATACAGAGCAGATGATACATATTAGAG 2355
QY 479 AGATGAGCTGCGCGCAAGTGGAAAGCCAAAGATGATCGCGCGCATCGCGCGCTTCATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGGAACCCAAATGATAGGGGGAATTGGAGTTTATCA 2415
QY 539 AGGTGCGCAGTACGACGAGATCTGATCGAGATCTGCGGCAAGAGCCCATCGGCACCG 598
Db 2416 AAGTAGGACAGTATGATCAGATCTCATAGAAATCTGCGGACATTAAGCTATAGTACAG 2475
QY 599 TGCTGATCGGCCCCACCCCGTGAACATCATCGCGCGCGCGCGCGCGCGCGCGCGCG 658
Db 2476 TATTAGTAGGACCTACACCTGTCACATTAATTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCATCAGCCCCATCGAGACCGTGCCTGGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTCCCATTAGTCTTATGAGACTGTACCAAGTAAATTTAAAGCCAGAA 2595
QY 719 TGGAGGCCCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAATCAAGGCCCTGACCG 778
Db 2596 TGGATGCCCAAAAGTTAAACAATGGCCATTGACAGAGGAAATAAAGCATTAGTAG 2655
QY 779 CCATCTCGGAGGAGTGAAGAGGAGGCAAGATCAACAAGATCGGCCCGCGCGAGACCCCT 838
Db 2656 AATTGTACAGAAATGGAAGAGGAAAGAAATTTCAAAATTTGGGCTTGAATCCAT 2715
QY 839 ACAACACCCCGTGTTCGCCCATCAAGAGAGGACAGCACCAAGTGGCGCAAGTGGTGG 898
Db 2716 ACAATACTCCAGTATTTCGCCATTAAGAGAAAGACAGTACTAATGAGAGAAATTAGTAG 2775

Qy	899	ACTTCCGGCAGCTGACACAGCCACCAGGACTTCTGGGAGGTGCACCTGGGCATCCCC	958
Db	2776	ATTTCAGAGAAGCTTATAGAGAACTCAAGATTTCTGGGAAGTTTCAATTAGGAATACCAC	2835
Qy	959	ACCCCGCCGCTGAAGAAGAAAGAGCGTGACCGTCTGGACGTGGGCGACGCCCTACT	1018
Db	2836	ATCCTGCAGGGTTAAACACAGAAAATCAGTAACAGTACTTGGATGTGGCGATGCATATT	2895
Qy	1019	TCAGCGTGGCCCTGGACGAGGACTTCCGCAAGTACACGCCCTTCACCATCCCGAGCATCA	1078
Db	2896	TTTCAGTCCCTTAGATAAAGACTTCAGGAAGTACTGCTATTTCACATACCTAGTATAA	2955
Qy	1079	ACAACGAGACCCCGCATCCGCTACCAGTACAGTGTCTGCCCCAGGGCTGGAAGGGCA	1138
Db	2956	ACAATGAGACACACAGGATTAGATATCAGTACATATGCTCTCCACAGGATGGAAGGAT	3015
Qy	1139	GCCCCAGCATTTCCAGACGACATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACC	1198
Db	3016	CACGACGATAATTCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC	3075
Qy	1199	CCGAGATCGTGATCTACCA- ----GGCCCCCTGTAGCTGGGCAGCGACTGGAGTCG	1252
Db	3076	CAGACATAGTCACTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAATAG	3135
Qy	1253	GCAGACGCGCGCAAGATCGAGGAGCTGCGCAAGCACTCTCGCTGGGGGTTCTACCA	1312
Db	3136	GGCAGCATAGACAAAAATAGAGGAACTGAGACAACATCTCTTGAGTGGGATTACCA	3195
Qy	1313	CCCCGCAAGAACACAGAGAGAGCCGCCCTTCTCTGCCAT-----CGAGCTGCACC	1366
Db	3196	CACGACAAAAAACATCAGAAAGAACCTCCATTCCTTTGGATGGGTATTGAACTCCATC	3255
Qy	1367	CCGACAAGTGGACCGTGCAGCCATCGAGCTGCCCGAAGGAGAGCTGGAACCGTGAACG	1426
Db	3256	CTGATAAATGGACAGTACAGCCTATAGTCTGCCAGAAAAGACAGCTGNACTGCAATG	3315
Qy	1427	ACATCCAGAAGCTGTGTGGCAAGCTGAACTGGGCCAGCCAGATGTACCCCGGCATCAAGG	1486
Db	3316	ACATACAGAAATTAGTGTGGAAAATTGAAATGGCGAAGTTCAGATTATTCAGGGATTAAAG	3375
Qy	1487	TGGCGCAGCTGTGCAAGCTGTCCGCGCGCCACAGGCCCTGCACCGACATGTGCCCCCTGA	1546
Db	3376	TAAGGCATTTATGTAACTCTTAGGGNACCAAAAGCACTAACAGAGTAGTACCAGTAA	3435
Qy	1547	CCGAGGAGCGCAGCTGGAGCTGGCCGAGAACCGCAGAGATCTCGCGAGCCGCTGCACG	1606
Db	3436	CAGAAGAAGCAGAGCTAGAACTGGCAGAAAAACAGGGAGATTCTAAAAGAACCGGTACATG	3495
Qy	1607	CGGTGTACTAGCACCCACAGAGGACCTGGTGGCCGAGATCCAGAGCAGGGGCCACGACC	1666
Db	3496	GAGTGTATTATGACCCTCAAAAGACTTAATAGAGAAATACAGAACGAGGGCAGGGCC	3555
Qy	1667	AGTGGACCTACAGATCTACCAGGAGCCCTTCAAGAACCTGAAAGACGGCAAGTACGCCA	1726
Db	3556	AATGGACATATCAATTTATCAGAGCCATTAAAAATCTGAAAACAGGAAAATATGCAA	3615
Qy	1727	AGATGGCGACCCGCCACACACAGACGTGAAGAGCTGACCGAGGCCGTGCAGAAAGATCG	1786
Db	3616	GAATGAAGGGTGGCCACACTAATGATGTGAACAAATTAAACAGAGCGAGTACAAAAAATAG	3675
Qy	1787	CCATGGAGAGCATGTGATCTGGGCAAGACCCCAAGTCTCCGCTGCCCATCCAGAAAGG	1846
Db	3676	CCACAGAAAGCATAGTAATATGGGNAAGACTCTTAAATTTAAATTTACCCATACAAAGG	3735
Qy	1847	AGACCTGGGAGACCTGGTGGACCGACTACTTGGCAGGCCACCTGGATCCCCGAGTGGGAGT	1906
Db	3736	AAACATGGGAAGCATGCTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGGGAGT	3795
Qy	1907	TCGTGAACACCCCCCTGGTGAAGCTGTGTACCAGCTGGAGAGAGGAGGCCCATCATCG	1966
Db	3796	TTGTCAATACCCCTCCCTAGTGAAGTATTATGTTACCATGTAGAGAAGAACCCCAATAATG	3855
Qy	1967	GCGCCGAGACTTTCTACGTGGAGCGCGCGCCCAACCGCGGACCAAGATCGGCAAGGCCG	2026

RESULT 9

US-09-309-572-15

US-09-309-372-13 : Sequence 15 Application US/09309572

; sequence 13, Application
: Ratont No 6410730

; Patent No. 6440730

; GENERAL INFORMATION:

; APPLICANT: Heinrich-Pette-Institut

; TITLE OF INVENTION: Retroviral hyb

FILE REFERENCE: P50489

: CURRENT APPLICATION NUMBER: US/

; CURRENT APPLICATION NUMBER: US/03/303,372
 : CURRENT FILING DATE: 1999-05-11

;; CURRENT FILING DATE: 1999-03-11
; EARIER APPLICATION NUMBER: DE

; EARLIER APPLICATION NUMB.
: EARLIER FILING DATE: 100

; EARLIER FILING DATE: 1998-1

; NUMBER OF SEQ

; SOFTWARE: Pat

; SEQ ID NO 1

LENGTH: 9709

LENGTH: 3703
TYPE: DNA

Query Match	A3 88	Score	1075 2	DB 4	Length	9709
-------------	-------	-------	--------	------	--------	------

Query Match

Best Local Similarity 67.7%; Pred. No. 5.1e-168;

07 14 TCCCCCACCACCCACCA - - CCCCACATCCAGCCACCAT 70

QY I4 TGGCCGAGGCCATGAGCCAGGCCACCA--GCGCCAAACATCCTGATGCAGCGCAGCAACT 70

The image shows a gel electrophoresis result with 18 lanes. Lanes 1-4 contain molecular weight markers. Lanes 5-18 show various DNA bands. Lanes 13-18 show prominent bands at approximately 1000 bp, which are likely the plasmid DNA from the *Agrobacterium* strains.

Db 1877 TGGCTGAAGCAATGAGCCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATT 1936

QV 71 TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTGGGCAAGAGGCCACATCGCCC 130

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

D_b 1937 TTAGGAAACCAAGAAAGACCTGTTAAGTGTTCATTGTGGCAAGAGAGGCACATAGCCA 1996

DB 1937 TTAGGAACCAAAGAAAGACTGTAAAGTGTTCATTTGTGGCAAAGAAGGGCACATAGCCA 1998

100

QY 131 GCAACTGCCGCGCCCCCGCAAGAAGGGCTGCTGGAAGTGCAGGAGGCCACCAGA 190


```
; TITLE OF INVENTION: LENTIVIRAL VECTORS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,312
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: CHANG-112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12494 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA"
; US-08-935-312-13

Query Match 43.8%; Score 1075.2; DB 4; Length 12494;
Best Local Similarity 67.7%; Pred. No. 5.2e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGCGCCGAGGCATGAGCCAGGCCACCA---GCCGCAACATCCTTGATGCGCGCAGCAACT 70
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1548 TGCGTGAAGCAATGAGCCAGTAAACAAATCCAGCTACCAATATGATACAGAAAGCAATT 1607

QY 71 TCAAGGCCGCCAGGATCATCAAGTGTTCAACTGCGGCAAGGAGGCGGCACATCGCCC 130
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1608 TTAGGAACCAAGAAAGAACTGTTAACTGTTCAATTGTGCAAGAAAGGSCACATAGCCA 1667

QY 131 GCAACTGCCGCGCCCGCCGCAAGAGGCTGCTGGAAGTGGCGCAAGGAGGCCACCGA 190
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1668 AAAATTGCAAGGCGCCCTAGGAAAGAGGCTGTTGGAATGTGGAAGAGGAGACACCAA 1727

QY 191 TGAAGGACTGCACCGGCGCGCAACTTCTTCGCGGAGGACCTGGCTTCCCGCAGG 250
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1728 TGAAGATGTACTGAGACAGAGCTAA- TTTTATTAGGAAGATCTGGGCTTCCCAAG 1786

QY 251 GCAAGGCCCGGAGTTCCTCCAGGAGCAGACAGCCGCAACGCCACCGCGCGGAGC 310
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1787 GGAAGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACAGCCGCCACAGAGAGAGC 1846

QY 311 TGCAGTGGCGG- - - - -CGACACCCCGCAGCGAGGCGCGCGCGGCGCAGGGCA 364
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1847 TTCAGGTTGGGGAAGAGACAACTCCCTCTCAGAAGCAGGAGCGGATAGACAAGGAA 1906

QY 365 - - - - -CCCTGAACCTTCCCGCAGATCACCTGTGGCAGCGCCCTGGTGAGCATCAAG 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1907 CTGTATCTTTAGCTTCCTTCAGATCACCTTTGGCAGGACCCCTCGTCAATAAAGA 1966

QY 419 TGGGGCGGCAGATCAAGGAGGCGCTGCTGGACACCGCGCGACACACCGTGTGAGG 478
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1967 TAGGGGGCAATTAAAGGAAGCTTATTAGATACAGGACGAGATGATACAGTATTAGAAG 2026

QY 479 AGATGAGCTGCCGCGCAAGTGAAGCCCAAGATGATCGGCGGATCGGCGGCTTCATCA 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31,677
REFERENCE/DOCKET NUMBER: 015280-249000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 15581 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: - 1..15581
LOCATION: 1..15581
OTHER INFORMATION: /note= "pNLSg11"

US-08-646-538-35

Query Match 43.8%; Score 1075.2; DB 3; Length 15581;
Best Local Similarity 67.7%; Pred. No. 5.3e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;

QY 14 TGCCGAGGCCATGAGCCAGGCCACCA---GCCGCCAACATCCTGTGATGCGAGCGCAGCAACT 70
DB 1877 TGGCTGAAGCAATGAGCCAGTAACAAATCCAGCTACCATAATGATACAGAAAGGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTTCAGCTGCGCGCAAGGAGGGCCACATCGCCC 130
DB 1937 TTAGGAACCAAGAAAGACTGTAAAGTGTTCATTTGTCGAAGAGAGGGCCACATAGCCA 1996
QY 131 GCAACTGCGCGCCCCCGCAAGAGGGCTGCTGGAAGTGCAGCAAGGAGGGCCACACAGA 190
DB 1997 AAAATGTCAGGGCCCCCTAGGAAGAGGGCTGTGGAAATGTGAAGAGGAGACACAAA 2056
QY 191 TGAAGACTGCAACGAGCGCCAGGCCAACTTCTTCGCGAGGACCTGGCCCTCCCGCCAGG 250
DB 2057 TGAAGATTGTACTGAGACAGAGCTAA-TTTTTTAGGGAAGATCTGGCCCTCCACAA 2115
QY 251 GCAAGGCCCGGATGTCAGGAGGAGCAGACAGCCGCGCAACAGCCGCCACCGCGGAGC 310
DB 2116 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCACAGCCGCCACAGAGAGAGC 2175
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCAGCGCCAGGGCA 364
DB 2176 TTCAGGTTGGGAAGAGACAACTCCCTCTCAGAGCAGGAGCGCGATAGACAGGAA 2235
QY 365 -----CCCTGAATTCGCCAGATCAACCTGTGGCAGCGCCCGCTGGTGAGCATCAAG 418
DB 2236 CTGTATCCTTTAGCTTCCCTCAGATCACCTCTTTGGCAGCGACCCCTCGTCACAATAAGA 2295
QY 419 TGGCGGCCAGATCAAGGAGGCCCTGCTGACACCGCGCGCGAGCACCGTGTGGAGG 478
DB 2296 TAGGGGGGCAATTAAGAGAGCTCTATATAGATACAGGAGCAGATGATACAGTATTAGAG 2355
QY 479 AGATGAGCCTGCGCGGCAAGTGAAGCCCAAGATGTCGCGGCATCGCGCGCTTCATCA 538
DB 2356 AATGAAATTTGCCAGNAGATGAAACCAAAATGATAGGGGAATTTGAGGTTTATCA 2415
QY 539 AGGTGCGCCAGTACGACCCAGATCCTGATCGAGATCTGCGCAAGAGGCCATCGGCACCG 598
DB 2416 AAGTAGGACAGTATGATCAGATCTCATAGAAATCTCGCGACATTAAGCTATAGGTACAG 2475
QY 599 TGTGATCGGCCCGCCCGTGAACATCATCGCGCCACACATGCTGACCCAGTGGGCT 658
DB 2476 TATTAGTAGGACCTTACACCTGTCAACATAATTTGGAAGAAATCTGTTGACTCAGATTGGCT 2535
QY 659 GCACCTGTAATTTCCCATCAGCCCCATGAGACCGTGGCCGTGAAGCTGAAGCCCGGCA 718
DB 2536 GCACTTTAAATTTTCCCATATTGCTCTATTGAGACTGTACCAGTAAATTTAAGCCAGGAA 2595
QY 719 TGGACGGCCCCAAGGTGAAGCAGTGGCCCTGACCGGAGGAGAGATCAAGGCCCTGACCG 778

DB 2596 TGGATGCCCAAAAGTTAAACAATGGCCATTGACAGAAGAAAAATAAAGCATTTAGTAG 2655
QY 779 CCATCTCGGAGGAGATGGAGAAGGCGAAGATCAACAAGATCGGCCCGGAGAACCCCT 838
DB 2656 AAATTTGTACAGAAATGGAAAAGAGGAAAAATTTCAAAAATTTGGGCTTGAANAATCAT 2715
QY 839 ACAACACCCCGCTGTTGCGCATCAAGAAGAGGACACACCAAGTGGCGCAAGTGGTGG 898
DB 2716 ACAATACTCCAGTATTTGCCATAAAGAAAAAGACACTACTAAATGAGAAAAATTAGTAG 2775
QY 899 ACTTCGCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGGTGCGAGCTGGGCATCCCC 958
DB 2776 ATTTTCAGAGAACTTAATAAGAGAACTCAAGATTTCTGGGAAGTTCAATTTAGGATACAC 2835
QY 959 ACCCGCGCGGCTGAAGAAGAAGAGCGTGCAGCGTGTGAGCGTGGGCGGCGACCCCTACT 1018
DB 2836 ATCTTCGAGGGTTAAACAGAAAAAATCAGTAACTAGTGGATGTGGGCGATGCATATT 2895
QY 1019 TCAGCGTGCCTCGGAGGAGACTTCGCAAGTACACCGCTTCACCATCCCCAGCATCA 1078
DB 2896 TTTTCAGTTCCCTTAGATAAAGACTTCAGGAAGTATACCTGATTTACCATACCTAGTATA 2955
QY 1079 ACAACGAGACCCCGGATCCGCTACCACTACAGTACACGCTGCTGCCCCAGGGTGGGAAGGCA 1138
DB 2956 ACAATGAGACACCCAGGATTTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAGGAT 3015
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCCGCGCCGCAACC 1198
DB 3016 CACGAGCAATATTTCCAGTGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAAATAATC 3075
QY 1199 CCGAGATCGTGATCTACCA-----GGCCCCCTGTGAGTGGGCGAGCAGCTGAGATCG 1252
DB 3076 CAGACATAGTCACTCATCAATACATGATGATTTGTATGTAGGATCTGACTTACAAATAG 3135
QY 1253 GCAGACACCGGCGCAAGATCGAGGAGCTGCGCAAGCAGCTGCTGCGTGGGGCTTCACCA 1312
DB 3136 GGCAGCATAGAAAAAATAGAGGAAGTGAACAACATCTGTTGAGGTGGGGATTTACCA 3195
QY 1313 CCGCCGACAGAGACACCAAGAGGAGCCCTTCCCTGCCCAT-----CGAGCTGCACC 1366
DB 3196 CACGACAAAAAACAATCAGAAAAGAACCTCCATTTCTTGGATGGGTATGAACTCCATC 3255
QY 1367 CCGACAAAGTGCAGCTGCAGCCCATCGAGTGCAGGAGAGAGAGAGCTGGACCTGGAACG 1426
DB 3256 CTGATAATGGACAGTACAGCTTATGCTGTCGCAAGAAAGGACAGCTGAGCTGCTAATG 3315
QY 1427 ACATCCAGAGAGCTGGTGGGCAAGCTGAACTGGGCGCAGCCAGATCTACCCGGCATCAAG 1486
DB 3316 ACATACAGAAATAGTGGGAAAAATGAATTTGGCAAGTCAAGATTTATGCAAGGATTAAG 3375
QY 1487 TGGCGCAGCTGTCAAGCTGCTGCGCGCGCCCAAGGCCCTGACCGACATCCTGCGCCCTGA 1546
DB 3376 TAAGGCAATTTATGTAAGTCTTTAGGGGAACCAAGCACTAACAGAGTAGTAGCACATAA 3435
QY 1547 CCGAGAGGCGGAGCTGGAGTGGCGGAGAACCGGAGATCCTCGCGAGCGCCCTGACG 1606
DB 3436 CAGNAGAGCAGAGTAGAAGTGCAGAAAACAGGGAGATCTTAAAGAACCCGCTACATG 3495
QY 1607 GCCTGTACTACGACCCCGCAAGGAGCCTGCTGGCGGAGATCCAGAAAGCGGGCCAGCACC 1666
DB 3496 GAGTGTATTATGACCCCATCAAAAGACTTAATAGCAGAAATACAGAAAGCAGGGCAAGGCC 3555
QY 1667 AGTGGACCTACCAGATCTACCAGGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA 1726
DB 3556 AATGGACATATCAAAATTTATCAGAGCCATTTAAAAATCTGANAACAGGAAATATGCAA 3615
QY 1727 AGATGCGACCGCCACACCAACGCTGAAGCAGCTGACCGAGCGCGTGCAGAAAGTCG 1786
DB 3616 GAATGAAGGTGCCCCACACTAATGATGTGAACAAATTAACAGAGCGCAGTACAAAAATAG 3675
QY 1787 CCATGGAGAGCATGCTGATCTGGGGCAAGCCCCCAAGTTCCGCTGCCCATCCAGAGG 1846

Db 3676 CCACAGAAAGCATAGTAAATATGGGAAAGAGCTCTAAATTTAAATTAACCATACAAAGG 3735
QY 1847 AGACCTGGAGACCTGGTGACCGACTACTGGCAGGCCACTGGATCCCGGAGTGGAGT 1906
Db 3736 AAACATGGGAAGCATGGTGACAGAGATTGGGAAGCCACTGGATTCTGAGTGGGAGT 3795
QY 1907 TCCTGAACACCCCGCCCTGGTGAAGCTGGTGGTACAGCTGGAGAGGCCATCATCG 1966
Db 3796 TTGTCAATACCCCTCCCTAGTGAAGTTAGTACAGCTGGAGAGAACCCATAATAG 3855
QY 1967 GCGCCGAGACCTTCTAGTGGAGCGGCCGCCCAACCCGAGACCAAGATCGGAAGCCG 2026
Db 3856 GAGCAGAAACTTTCTATGTAGATGGGCGAGCCCAATAGGGAATACTAAATTTAGGAAGACG 3915
QY 2027 GCTACGTGACCGCGCGCGGCGGAGAGATCGTGACCTTGACCGAGACCAACCAACAGA 2086
Db 3916 GATATGTAACCTGACAGAGGAAGCAAAAGTTGTCCCGCTTAACGGACACCAACAAATCAGA 3975
QY 2087 AGACCGAGCTGCAGGCGCATCCAGCTGGCCCTGCAGGAGACGCGGAGGTGAACATCG 2146
Db 3976 AGACTGATTTACAAGCAATTCATCTAGCTTGCAGGATTCCGGATTAGAAGTAACATAG 4035
QY 2147 TGACCGACACGAGTAGTGGCCCTGGGCATATCCAGGCCCGCCGACCAAGAGGAGAGCG 2206
Db 4036 TGACAGACTCACAAATATGCAATGGGAATCATTTCAAGCACCAACAGATAAGAGTGAATCAG 4095
QY 2207 AGCTGGTGAACCATCATCGACGAGCTGATCAAGAGGAGAGAGGTGTACTGAGCTGGG 2266
Db 4096 AGTTAGTCACTCAANTATAGACGAGTTAATAAAGGAAAGTCTACTGGCATGGG 4155
QY 2267 TGCCCGCCCAACAGGCGATCGCGGCAACGAGCAGATCGACAAGCTGGTGAGCAAGGCCA 2326
Db 4156 TACCAGCACACAAAGGAATTTGGAGGAATCAACAAGTAGATGGTGGTCAAGTCTGCTGGA 4215
QY 2327 TCAGCAGAGTGGTCTTCTCGGAGCGGATCGAT 2358
Db 4216 TCAGGAAAGTACTATTTTAGATGGAATAGAT 4247

RESULT 13

US-09-503-222-35
; Sequence 35, Application US/09503222
; Patent No. 6265548
; GENERAL INFORMATION:
; APPLICANT: Pavlakis, George N.
; APPLICANT: Gaitanaris, George A.
; APPLICANT: Stauber, Roland H.
; APPLICANT: Vournakis, John N.
; TITLE OF INVENTION: Mutant Aequorea victoria Fluorescent
; TITLE OF INVENTION: Proteins Having Increased Cellular Fluorescence
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09503,222
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,538
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Kenneth A.
; REGISTRATION NUMBER: 31,677

; REFERENCE/DOCKET NUMBER: 015280-249000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: - 1..15581
; LOCATION: 1..15581
; OTHER INFORMATION: /note= "pNlnsg11"
US-09-503-222-35

Query Match 43.8%; Score 1075.2; DB 4; Length 15581;
Best Local Similarity 67.7%; Pred. No. 5 3e-168;
Matches 1606; Conservative 0; Mismatches 738; Indels 28; Gaps 6;
QY 14 TGGCCGAGCCATGAGCCAGGCCACCA---GCGCCACATCTCTGATCGAGCGCAGCAACT 70
Db 1877 TGGCTGAAGCAATGAGCAAGTAACAAATCCAGCTACCATAATGATACAGAAAGCAATT 1936
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGCTTCAACTCGGCAAGGAGGCCACATCGCCC 130
Db 1937 TTAGGAACCAAGAAAGACTGTTAAGTGTTCATTTGTCGCAAGAGGGCACATAGCCA 1996
QY 131 GCAACTGCCCGCGCCCGCCGCAAGAGGCTGTGGAAGTGGCGCAAGGAGGCCACCA 190
Db 1997 AAATTCAGCGGCCCTAGGAAGAGGGCTGTGGAATGTGGAAGAGGAGGACACCAA 2056
QY 191 TGAAGGACTCGACGCGCGCCAGCCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCCCAGG 250
Db 2057 TGAAGATTTACTTGAGAGACAGGCTAA-TTTTATTAGGAAGATCTGGCCTTCCCAACG 2115
QY 251 GCAAGGCCCGCGAGTTCCCGCAGGAGCAAGCCGCGCAACAGCCGCCAGCGCCGCGAGC 310
Db 2116 GGAAGGCCAGGGAATTTCTTCAGAGCAGACAGAGCCAGAGCCGCCAGCAAGAGAGC 2175
QY 311 TGCAGGTGCCGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCCAGGGCA 364
Db 2176 TTCAGGTTTGGGAAGAGAGACAACACTCCCTCTCAGAAGCAGGAGCGCGATAGACAAGAA 2235
QY 365 -----CCCTGAACCTCCCGCCAGATCACCTGTGGCAGCGCCCGCTGGTGAGCATCAAGG 418
Db 2236 CTGTATCTCTTTAGCTTCCCTCAGATCACTCTTTGGCAGCGACCCCTCGTCACATAAAGA 2295
QY 419 TGGCGCGCCAGATCAAGAGGCGCTGTGTCACACCGCGCGCGAGCACCGCTGTGAGG 478
Db 2296 TAGGGGGCAATTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGATTAGAAG 2355
QY 479 AGATGAGCCTGCCCGCAAGTGAAGCCCAAGATGATCGCGCGCATCGCGCGCTTCAATCA 538
Db 2356 AAATGAATTTGCCAGGAAGATGAAACCAAAATATGAGGGGAATTTGAGGTTTTATCA 2415
QY 539 AGGTGCGCCAGTAGCAGCCAGATCCTGATCGAGATCTCGGCAAGAGGCCATCGGCACCG 598
Db 2416 AGTAGGACAGTATGATCATAGATACTCATAGAAATCTCGGACACATAAAGCTATAGGTACAG 2475
QY 599 TGTGTATGGGCGCCCGCCCGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT 658
Db 2476 TATTAGTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTGACACAGATTGGCT 2535
QY 659 GCACCTGAACTTCCCCATCAGCCCCCATCGAGACCGCTGCCCGTGAAGCTGAAGCCCGCA 718
Db 2536 GCACCTTAAATTTTCCCATTTAGTCTTATTGAGACTGTACCACTGTAAGAGCCAGAA 2595
QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTGACCGAGGAGAGATCAAGGCCCTGACCG 778
Db 2596 TGGATGGCCCAAGGTTAAACAATGGCCCTTGACAGAAAGAAAAATAAAGCATATTAGTAG 2655

Qy	779	CCATCTCGAGGAGATGGAGAGGAGGGCAAGATCACCAAGATCGGCCCGCAGAACCCCT	838
Db	2656	AAATTTGTACAGAAATGGAAAAGAGGAAAAAATTTCAAAAATTTGGGCTCGAAAATCCAT	2715
Qy	839	ACAAACCCCGTGTTCGCATCAAGAAAGAGGACAGCAACCAAGTGGCGAAGCTGTGG	898
Db	2716	ACAATCTCCAGTATTGGCATTAAGAAAAAGACAGTACTAAATGGAGAAAAATTTAGTAG	2775
Qy	899	ACTTCCCGAGCTGAACAAGCGCACCCAGGACTTCTGGGAGTGTGAGCTGGGCATCCCCC	958
Db	2776	ATTTACAGAACTTAATAAGAGAACTCAAGATTTCTGGGAGTTCAATTAGAAATACAC	2835
Qy	959	ACCCCGCGGCTGAAGAAAGAGAGAGCGGTGACCGTGTGGAGTGGGCGACGCCCTACT	1018
Db	2836	ATCCTGCAAGGTTTAAACACAGAAAAATCAGTAACAGTACTTGATGTGGCGATGCATTT	2895
Qy	1019	TCAGCGTCCCTGGAGGAGGACTTCCGCAAGTACACCGCTTACCATCCCCAGCATCA	1078
Db	2896	TTTCAGTTCCTTAGATAAAGACTTCAGGAAGTATACTGCAATTTACCATCTAGTATAA	2955
Qy	1079	ACAACGAGACCCCGGCATCCGCTACCACTACAACGCTGCTGCCACAGGCTGGAAGGCA	1138
Db	2956	ACAATGAGACACCGAGTATAGATATCACTAGTACAATGCTCTCCACAGGATGGAAAGAT	3015
Qy	1139	GCCCCAGCATTTCCAGACAGATGACCAAGATCCTCGAGCCCTTCGCGCCGCAACC	1198
Db	3016	CACGAGCAATTTCCAGTGTAGCATGACAAAAATCTTAGAGCTTTTAGAAAAACAAATC	3075
Qy	1199	CCGAGATCGTATCTACCA- - - - -GGCCCCCTGTAGTGGGAGCGACCTGGAGATCG	1252
Db	3076	CAGCATAGTCTATCTAATACATGGATGATTTGTATGTAGTATCTGACTTAGAAATAG	3135
Qy	1253	GCCAGCACCGCCAAAGATCGAGGAGCTGCGCAAGCACTGCTGCGTGGGGCTTTACCA	1312
Db	3136	GGCAGCATAGAACAANAATAGAGGACTGAGACACATCTGTTGAGGTGGGATTTACCA	3195
Qy	1313	CCCCGCAAGAAGACACAGAAGGAGCCCCCTTCTGCCCAT- - - - -CGAGCTGCACC	1366
Db	3196	CACGACAAAAAACATCAGAAGAACCCTCCATTCTTTGGATGGGTATTAAGCTCCATC	3255
Qy	1367	CCGACAACTGACCGTGCACCCATCGAGCTGCCCGAAGGAGAGCTGACCTGGAACG	1426
Db	3256	CTGATAAATGGACAGTACAGCTATAGTGTGCCAGAAAGGACAGCTGGAGTCTCAATG	3315
Qy	1427	ACATCCAGAGCTGTGTGGCAAGCTGAATGGGCCAGCCAGATCTACCCGGCATCAAGG	1486
Db	3316	ACATACAGAAATTAGTGGGAAATTAATTTGGCAAGTCAGATTTATGCAAGGATTAAG	3375
Qy	1487	TGCGCCAGCTGTCAAGCTGTCTGCGGGGGCAAGGCCCTGACCGACATCTGTGCCCTGA	1546
Db	3376	TAAGGCAATTTATCTAACTTCTTAGGGAAACCAAGCACTAACAGAAGTAGTACCACTAA	3435
Qy	1547	CCGAGGAGCCGAGCTGGAGCTTGGCCGAGAACCCGAGATCCTGCGGAGCCCGTGAACG	1606
Db	3436	CAGAAGAGCAGAGCTAGAACTTGGCAGAAACAGGGAGATTTCTAAAAGAACCGGTATAG	3495
Qy	1607	GCGTGTACTACGACCCACCAAGGACCTGTGTGGCGAGATCCAGAAGCGGCGCACGACC	1666
Db	3496	GAGTGTATTATGACCATCAAAAGACTTAATAGCAGAAATACAGACAGGGCGCAGGCC	3555
Qy	1667	AGTGGACCTACCATGATCTACCAAGAGCCCTTCAAGAACCTTGAAGACCGGCAAGTACGCCA	1726
Db	3556	AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTTGAAAACAGGAAATATGCA	3615
Qy	1727	AGATGCGCACCCACACCAAGAGCTGAAGCAGCTGACCGAGGCGGTGAGAAATCG	1786
Db	3616	GAATGAAGGGTGCCACACTAATGATGTGAACAAATTAACAGAGGCAAGTACAAAAATAG	3675
Qy	1787	CCATGGAGAGCATCGTATCTGGGCAAGACCCCAAGTTCGCGCTGCCATCCAGAGG	1846
Db	3676	CCACAGAAAGCATAGTAATATGGGGAAGAGCTCTTAAATTTAAATTAACCATACAAAGG	3735
Qy	1847	AGACCTGGGAGACCTGGTGGACCGCACTACTGCGAGGCCACCTGGATCCCCGAGTGGGAGT	1906

Db 3736 AAACATGGGAGCAATGTTGACAGAGATTTGGCAAGCCACCTGGGATTCCTGAGTGGGAGT 3795
 Qy 1907 TCGTGAACACCCCCCCTGGTGAAGCTGTGGTACCAGCTGGAGAGGAGGCCATCATCG 1966
 Db 3796 TTGTCAATACCCCTCCCTTAGTGAAGTTATGGTACCAGTTAGAGAAAGACCCATATAG 3855
 Qy 1967 GCGCCGAGACCTTCTAGCTGGAGCGGCCGCCAACCGCGAGACCAAGATCGGCAAGCGC 2026
 Db 3856 GAGCAGAAACTTTCTATGTAGATGGGGCAGCCAATAAGGAAAACCTAAATTAGGAAAAGCAG 3915
 Qy 2027 GCTACGTGACCGCGGGCCGCGAGAAGATCGTGAGCCTGACCGAGACCCACCAACCCAGA 2086
 Db 3916 GATATGTAAGTACGACGAGGAAGACAAAAGTTGTCCCCCTAACGGACACACAATCAGA 3975
 Qy 2087 AGACCGAGCTGCAGGCCATCCACTGGCCCTGCAGGACACGCGGACGAGGTGAACATCG 2146
 Db 3976 AGACTGAGTTACAAGCAATTCATCTAGCTTTCAGGATTCGGGATTAGAAGTAAACATAG 4035
 Qy 2147 TGACCGACACCCAGTAGCCCTGGGCATCATCCAGGCCACGCCGACGAGCGAGCG 2206
 Db 4036 TGACAGACTCACAAATATGCAATGGGAATCATTCACGACACCAACCGATGAAGTGAATCAG 4095
 Qy 2207 AGCTGGTGAACCAAGATCATCGACGAGCTGATCAAGAAGGAGGAAGTGTACCTGAGCTGGG 2266
 Db 4096 AGTTAGTCAGTCAATTAATAGACAGCTTAATAAAAAGGAAAAAGTCTACCTGGCATGGG 4155
 Qy 2267 TGCCCCGCCACAAAGGGCATCGCGCGCAACGACGACATCGACAAGCTGTGTGACGAAGGCA 2326
 Db 4156 TACCAGCACACAAAGGAATTTGGAGGAATGAACAAGTAGATGGTTGGTCAAGTCTGCTGAA 4215
 Qy 2327 TCCGCAAGGTGCTCTTCCCTGGACGGCATCGAT 2358
 Db 4216 TCAGGAAGTACTATTTTAGATGGAAATAGAT 4247

 RESULT 14
 US-08-944-449-7
 ; Sequence 7, Application US/08944449
 ; Patent No. 5985613
 ; GENERAL INFORMATION:
 ; APPLICANT: KURTH, REINHARD
 ; APPLICANT: BAIER, MICHAEL
 ; APPLICANT: METZNER, KARIN
 ; APPLICANT: WERNER, ALBRECHT
 ; TITLE OF INVENTION: Use of an "immunodeficiency-virus suppressing
 ; TITLE OF INVENTION: lymphokine (ISL)" to inhibit the replication of
 ; TITLE OF INVENTION: viruses, particularly of retroviruses
 ; FILE REFERENCE: 8341-7065
 ; CURRENT APPLICATION NUMBER: US/08/944, 449
 ; CURRENT FILING DATE: 1997-10-06
 ; EARLIER APPLICATION NUMBER: EP 95113013.2
 ; EARLIER FILING DATE: 1995-08-18
 ; EARLIER APPLICATION NUMBER: DE 195 13 152.5
 ; EARLIER FILING DATE: 1995-04-07
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 9737
 ; TYPE: DNA
 ; ORGANISM: Human immunodeficiency virus type 1
 US-08-944-449-7

[illegible]

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:45:48 ; Search time 107.404 Seconds
(without alignments)
10843.369 Million cell updates/sec

Title: US-09-610-313-32

Perfect score: 2457

Sequence: 1 gtcgagccaccatggccga.....gggctagcaccgggaattc 2457

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US50_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US50_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1607.4	65.4	4307	10	US-09-999-183-2
2	1539.2	62.6	8366	10	US-09-872-733-6
3	1503.8	61.2	4338	10	US-09-872-733-1
4	1137	46.3	2467	10	US-09-872-733-3
5	1081.6	44.0	2601	10	US-09-735-487-7
6	1081.6	44.0	2601	10	US-09-735-487-9
7	1081.6	44.0	2601	10	US-09-735-487-11
8	1081.6	44.0	2601	10	US-09-735-487-13
9	1081.6	44.0	4307	10	US-09-999-183-1
10	1081.6	44.0	9719	10	US-09-737-190A-1
11	1029.2	41.9	8933	10	US-09-943-286-3
12	1029.2	41.9	8933	10	US-09-943-286-4
13	1029.2	41.9	8933	10	US-09-943-286-9
14	1022.8	41.6	9544	10	US-09-798-675-4
15	1022.8	41.6	9544	10	US-09-798-675-5
16	818.4	33.3	2348	9	US-10-097-997-1
17	805.8	32.8	9793	9	US-09-886-156-56
18	805.8	32.8	9793	9	US-09-886-150-56
19	805.8	32.8	9793	9	US-09-886-149-56

20 805.8 32.8 9793 9 US-09-886-159-56 Sequence 56, Appl
21 699.8 28.5 2507 10 US-09-872-733-2 Sequence 2, Appl
22 657.2 26.7 12379 10 US-09-991-258-14 Sequence 14, Appl
23 654 26.6 1323 10 US-09-991-258-15 Sequence 15, Appl
24 425.4 17.3 948 10 US-09-756-551A-5 Sequence 5, Appl
25 243.8 9.9 1800 9 US-10-003-035-58 Sequence 58, Appl
26 239.2 9.7 4516 10 US-09-968-355-24 Sequence 24, Appl
27 216.4 8.8 12523 10 US-09-991-258-1 Sequence 1, Appl
28 215.6 8.8 1476 10 US-09-991-258-4 Sequence 4, Appl
29 172.6 7.0 507 9 US-10-003-035-56 Sequence 56, Appl
30 163.2 6.6 1482 10 US-09-818-443-4 Sequence 4, Appl
31 163.2 6.6 1532 10 US-09-818-443-1 Sequence 1, Appl
32 157.8 6.4 1503 10 US-09-968-355-25 Sequence 25, Appl
33 156.6 6.4 1752 10 US-09-968-355-16 Sequence 16, Appl
34 156.6 6.4 4594 10 US-09-968-355-15 Sequence 15, Appl
35 138.2 5.6 9464 10 US-09-738-847-1 Sequence 1, Appl
36 133.6 5.4 1496 9 US-10-003-035-17 Sequence 17, Appl
37 130.6 5.3 13049 9 US-09-996-073-1 Sequence 1, Appl
38 130.6 5.3 13049 10 US-09-231-235-1 Sequence 1, Appl
39 130.6 5.3 13049 10 US-09-797-518A-1 Sequence 21, Appl
40 112.6 4.6 2310 10 US-09-476-242-21 Sequence 13, Appl
41 112.6 4.6 2535 10 US-09-476-242-13 Sequence 6, Appl
42 111 4.5 2328 10 US-09-476-242-6 Sequence 15, Appl
43 111 4.5 2523 10 US-09-476-242-15 Sequence 16, Appl
44 108.6 4.4 2517 10 US-09-476-242-16 Sequence 17, Appl
45 108.6 4.4 2517 10 US-09-476-242-17 Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-999-183-2
; Sequence 2, Application US/09999183
; Patent No. US20020147169A1
; GENERAL INFORMATION:
; APPLICANT: MITROPHANOUS, et al
; TITLE OF INVENTION: In Vivo Selection Method
; FILE REFERENCE: 674523-2009
; CURRENT APPLICATION NUMBER: US/09/999,183
; CURRENT FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: PCT/GB00/02136
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 9912965.2
; PRIOR FILING DATE: 1999-06-03
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: SeqWin99
; SEQ ID NO 2
; LENGTH: 4307
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Codon optimised gagpol sequence
US-09-999-183-2

Query Match 65.4%; Score 1607.4; DB 10; Length 4307;
Best Local Similarity 81.7%; Pred. No. 2.4e-256;
Matches 1939; Conservative 0; Mismatches 406; Indels 28; Gaps 6;
QY 12 CATGCCGAGGCCATGAGCCAGG---CCACCAGCGCCCAACCTCTGATGTCAGCGCCAGCAA 68
1086 CTGCTGTCGAGCCATGAGCCAGGTGACCACTCCGCTACCATCATGATGTCAGCGCCGCAA 1145
DB 69 CTTCAAGGGCCCAAGCGCATCATCAAGTGTCAACTGCGGCAAGGAGGCCACATGCG 128
1146 CTTTCGGAACCAACCAAGCAAGTCGTCAAGTGTCAACTGTGCAAGAAAGGACACAGC 1205
QY 129 CGCAACTGCCCGCCCGCCCGCCCGCTGCTGGAAGTGGCCCAAGGAGGCCACCA 188
1206 CGCAACTGCCCGCCCGCCCGCTGCTGGAAGTGGCCCAAGGAGGCCACCA 1265
QY 189 GATGAAGGACTGCACCGCGCCAGCGCCCAACTTCTTCGCGGAGGACCTGGCTTCCCCCA 248

Db 1266 GATGAAGACTGTACTGAGAGACAGGCTAA- TTTTATAGGAAGATCTGGCCCTCTCTACA 1324
QY 249 GGGCAAGGCCGCGAGTTCCCCAGCAGACAGAACCGCGCAACAGCCCCACACGCGCGCA 308
Db 1325 AGGAAGGCCAGGGAATTTCTTCAGAGCAGACACAGACCAACAGCCCCACAGAGAGA 1384
QY 309 GCTGAGGTGCGGG- - - - - CGACAAACCCCGCAGGAGCGCGCGCGAGCGCAGGG 362
Db 1385 GCTTCAGGTGTGGGTGAGACAACTCCCTCAGAGCAGGAGCCGATAGAAAG 1444
QY 363 CA- - - - - CCTCAACTTCCCCCAGATCACCTTGGCAGCGCCCTGGTGAAGATCAA 416
Db 1445 AACTGATCTTTAACTTCCTCAGATCACTTTGGCAACGACCCCTGCTCAATAAA 1504
QY 417 GGTGGCGGCAGATCAAGAGGCCCTGCTGGACACCGCGCGCAGACACCGTGTGGA 476
Db 1505 GATAGGGGCGACTCAAGAGGCTCTCTGGACACCGGAGCAGACACCGTGTGGA 1564
QY 477 GGAGATGAGCCTGCCCGCAAGTGGAAAGCCCAAGATGATCGGCGGATCGGCGCTTCAT 536
Db 1565 GGAGATGCTGTTGCCAGGCGCTGGAAAGCCGAAGATGATCGGGGAATCGGCGTTTCAT 1624
QY 537 CAAGGTGCGCAGTACGACAGATCCTGATCGAGATCTCGGCAAGAGCCATCGGCAC 596
Db 1625 CAAGGTGCGCAGTATGACAGATCCTCATCGAAATCTGGCGCCCAAGGCTATCGGTAC 1684
QY 597 CGTGCTGATCGGCCACCCCGTGAACATCATCGGCGCGCAACATGCTGACCCAGCTGG 656
Db 1685 CGTGCTGGTGGCCACACCCGCTCAACATCATCGGACGAACTGTTGACGCAGATCGG 1744
QY 657 CTCACACCCCGTGTGCCATCAAGCCCATCGAGACCGTGCCTGAAAGTGAAGCCCGG 716
Db 1745 TTCACGCTGAACCTCCCATAGCCCTATCGAGACGGTACCGGTGAAAGTGAAGCCCGG 1804
QY 717 CATGGAGGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAAGATCAAGCCCTGAC 776
Db 1805 GATGGAGGCCCGAAGGTCAAGCAATGGCCATTGACAGAGGAAGATCAAGGCACTGGT 1864
QY 777 CGCATCTGCGAGGATGAGAAAGAGGCAAGATCACCAAGATCGGCCCGCAGAACCC 836
Db 1865 GGAGATTTGCACAGAGATGAAAGAGGAAAGGAATCTCCAAGATTGGGCTGAGAACCC 1924
QY 837 CTACACACCCCGTGTGCCATCAAGAAAGAGACAGACCAAGTGGCGCAAGCTGGT 896
Db 1925 GTACACACCCCGGTGTGCCAATCAAGAAAGAGACTCGACAAATGGCGCAAGCTGGT 1984
QY 897 GGACTTCCGGAGCTGNACAGCGCACCCAGACTTCTGGGAGGTGCACTGGGCATCC 956
Db 1985 GGACTTCCGGAGCTGAACAGCGCACGCAAGACTTCTGGGAGGTTCAGCTGGGCATCC 2044
QY 957 CCACCCCGCGGCTGAAGAAGAAGAGCGTGACCGTCTGGACGTGGGCGACGCCTA 1016
Db 2045 GCACCCGAGGCTGAGNAGAAAGAAATCCGTGACCGTACTGGATGGGTGATGCCTA 2104
QY 1017 CTTTCAGGTGCCCTGGAGAGGACTTCGCAAGTACACCGCTTCACCATCCCGCAGCAT 1076
Db 2105 CTTCTCCGTTCCCTGGACGAACTTCAGGAAGTACACTGCCTTCACAATCCCTTCGAT 2164
QY 1077 CAACACGAGACCCCGCATCCGCTACAGATACAGTGTCTGCCCGCAGGGCTGGAAGG 1136
Db 2165 CAACACGAGACCCCGGATTCGATATCAGTACAACTGCTGCCCGCAGGGCTGGAAGG 2224
QY 1137 CAGCCCGCAGCATCTTCAGAGCAGCATGACCAAGATCTTGAGCCCTTCGCGCCCGCAA 1196
Db 2225 CTCCTCCGCAATCTTCAGATGAGATGACCAAAATCTTGAGCCCTTCGCGAAACAGAA 2284
QY 1197 CCCCAGATCGTGTACCA- - - - - GGCCCCCTGTAGTGGGAGCGACCTGGAGAT 1250
Db 2285 CCCCAGATCGTGTATCTATCAGTACATGGATGACTTGTAGTGGGCTGCTATCTAGAT 2344
QY 1251 CGGCCAGCACCAGCATCGAGAGCTGCGCAGACCTGCTGGCTGGGCTTCAC 1310
Db 2345 AGGGCAGCACCCAGCATCGAGAGCTGGCCAGACCTGTTGAGGTGGGACTGAC 2404

QY 1311 CACCCCGCACAAGAAGCACCAGAGGAGCCCTTCTCTGCCAT- - - - - CGAGCTGCA 1364
Db 2405 CACACCCGACAAGAAGCACCAGAGGAGCCCTCTCTCTGGATGGTTACGAGCTGCA 2464
QY 1365 CCCCAGCAAGTGAACCGTGCAGCCCATCGAGCTGCCCGAAGAGAGCTGGACCGTGAA 1424
Db 2465 CCGTGCAAAATGGACCGTGCAGCTTATCGTCTGCCAGAGAAAGACAGCTGGAGTGCAA 2524
QY 1425 CGACATCCAGAAGCTGGTGGCAAGCTGAATGGCCAGCCAGATCTACCCCGCATCAA 1484
Db 2525 CGACATACAGAAGCTGGTGGGAAGTTGAATGGGCCAGTCAAGTTTACCACAGGATTA 2584
QY 1485 GGTGGCGCAGCTGTGAAGCTGTGCGCGCGCAAGCCCTGACCCACATCTGCTGCCCT 1544
Db 2585 GGTGAGCAGCTGTGCAAACTCTCCGCGGAACCAAGGCACTCACAGAGTGATCCCT 2644
QY 1545 GACCGAGGAGCGAGCTGGAGTGGCGGAGAACCGCGAGATCTTGCAGAGCCCGTGCA 1604
Db 2645 AACCGAGGAGCGAGCTCGAATGCGAGAAACCCAGAGATCTTAAAGGAGCCCGTGCA 2704
QY 1605 CGCGTGTACTAGACCCCAAGCAAGGACCTGGTGGCGGAGATCCAGAGCAGGCGCACGA 1664
Db 2705 CGCGTGTACTATGACCCCTCCAAAGGACCTGATCGCGGAGATCCAGAGCAGGCGCAAGG 2764
QY 1665 CCAGTGGACCTACAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGC 1724
Db 2765 CCAGTGGACCTATCAGATTTACAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGC 2824
QY 1725 CAAGATGCGCACCGCCACACCAACGACAGTGAAGAGCTGACCGAGCCCTGCAAGAAGT 1784
Db 2825 CCGGATGAGGGGTGCCACACTAACGACGTCAAGAGCTGACCGAGCCCGTGCAAGAAGT 2884
QY 1785 CGCATGGAGAGCATCGTGTCTGGGCAAGACCCCAAGTTCGGCTGCCATCCAGAA 1844
Db 2885 CACCCAGAAAGCATCGTGTCTGGGAAAGACTCTTAAGTTCAAGTGGCCCATCCAGAA 2944
QY 1845 GGAGACTGGGACACCTGTGTGGACCACTACTGGCAGGCCACCTGGATCCCGAGTGGGA 1904
Db 2945 GGAACCTGGAAACCTGTGTGGACAGATATTGGCAGGCCACCTGGATCTCTGAGTGGGA 3004
QY 1905 GTTTCGTGAACACACCCCGCTGTGTGAAGTGTGTACAGTGGAGAAAGAGCCCATCAT 1964
Db 3005 GTTTCGTGAACACCCCTCCCTGTGTGAAGTGTGTACAGTGGAGAAAGAGCCCATAGT 3064
QY 1965 CGCGCGGAGACCTTCTACGTGGACGCGCCGCAACCGGAGACCAAGATTCGCAAGGC 2024
Db 3065 GCGCGCGAAACCTTCTACGTGGATGGGCGCGCTAACAGGGAGACTAAGTGGGCAAGC 3124
QY 2025 CGGCTAGCTACCGACCGCGGCGGAGAGATCGTGAGCTGACCGAGACCAACCAACCA 2084
Db 3125 CGATAGCTCACTAACCGGGGAGACAGAGTGTGTCACCTCACTGACACCAACCAACCA 3184
QY 2085 GAAGACCGAGCTCAGGCCATCCAGCTGGCCCTGCAAGCAGCGGAGAGGTGAACAT 2144
Db 3185 GAAGACTGAGCTCAGGCCATTTACCTGCTGCTGTCAGGACTCGGCGCTGGAGGTGAACAT 3244
QY 2145 CGTGACCGACAGCCAGTACGCCCTGGCATCATCCAGGCCCGCCCGCAGACGAGAG 2204
Db 3245 CGTGACAGACTCTCAGTATGCCCTGGCATCATTCGAAGCCCGCAGCAGAGTGAATC 3304
QY 2205 CGAGCTGGTGAACACAGATCATCGAGCAGCTGATCAAGAAGGAAAGGTGTACCTGAGCTG 2264
Db 3305 CGAGCTGGTCAATCAGATCATCGAGCAGCTGATCAAGAAGAAAGGTCTATCTGGCCTG 3364
QY 2265 GGTGCCCGCCCAAGGGCATCGCGGCAACGAGCAGATCGACAAGCTGTGTGAGCAAGG 2324
Db 3365 GGTACCCCGCCCAAAAGGCATTTGGCGCAATGAGCAGTTCGACAAGTGTGTCTCGGCTG 3424
QY 2325 CATCCGCAAGGTGCTTCTTGACGCGCATCGA 2357
Db 3425 CATCAGAAAGGTGCTATTCTCTGGATGGCATCGA 3457

RESULT 2

US-09-872-733-6
; Sequence 6, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 6
; LENGTH: 8366
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of the construct pCMVgagpolBTKan containing a CMV
; OTHER INFORMATION: promoter, a HIV gag/pol gene and a kanamycin
; OTHER INFORMATION: resistance gene
US-09-872-733-6

Query Match 62.6%; Score 1539.2; DB 10; Length 8366;
Best Local Similarity 79.9%; Pred. No. 4.1e-245;
Matches 1896; Conservative 0; Mismatches 448; Indels 28; Gaps 6;

QY 14 TGGCGGAGGCATGAGCCAGCCACAGC---GCCACATCCTGATGACGCGGAGCAACT 70
DB 1857 TGGCGGAGGCATGAGCCAGGCAGCAACTCGGCGACCATAAATGTCAGAGGCAACT 1916
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGCCTTCAACTCGGCGAAGGGGCCACATCGCCC 130
DB 1917 TCCGGAACCCAGCGAAGATCGTCAAGTGTCTCAATTGTGGCAAGAGGCCACCGCCA 1976
QY 131 GCAACTGCGCGGCCCGCCGCAAGAGGGCTGCTGGAAGTGCAGAGGAGGGCCACAGA 190
DB 1977 GGAATCGCGGGCCCCCGGAAGAGGGCTGTTGGAAATGGAAGAGGACACACAAA 2036
QY 191 TGAAGGACTGCACCGAGCGCCAGCCAACTTCTTCGCGAGGAGCTTGGCCCTTCCCCAGG 250
DB 2037 TGAAGATTGTCTGAGAGACAGGCTAA-TTTTATGGAAGATCTGGCCCTTCTTCAAG 2095
QY 251 GCAAGGCCGCGAGTTCCCGAGCGAGAGAACCGCGCCAAACAGCCCGCCAGCGCGGAGC 310
DB 2096 GGAAGGCCAGGGAATTTCTTTCAGAGCAGACAGAGCCACAGCCCGCCAGAGAGAGC 2155
QY 311 TCGAGGTGCGCGG-----CGACAAACCCCGAGCGAGGCGGCGCGGCGCGAGGCA 364
DB 2156 TTCAGGTCTGGGGTAGAGACAACAACCTCCCGCTCAGAAAGCAGGAGCGGATAGACAAGAA 2215
QY 365 -----CCCTGAATCTCCCGAGATACCTGTGGCAGCGCCCGCTTGGTGAGCATCAAGG 418
DB 2216 CTGTATCTTTAATCTCCCTCAGATCACTCTTTGGCAACGACCCCTCGTCACAGTAAGGA 2275
QY 419 TGGCGGCGCAGATCAAGGAGGCCCTGCTGGACACCGCGCGCCGACACACCGTGTGGAGG 478
DB 2276 TCGGGGGGCAACTCAAGGAAGCGCTGCTCGATACAGGAGCAGATGATACAGTATTAGAAG 2335
QY 479 AGATGAGCTGCGCGGCAAGTGGAGCCCAAGATGATCGCGGCGATCGCGGGCTTCAATCA 538
DB 2336 AAATGAGTTTGGCAGGAAGATGGAACCAAAAATGATAGGGGGGATCGGGGGCTTCAATCA 2395
QY 539 AGGTGCGCAGTAGCAGCATCTGATCGAGATCTCGGCAAGAGGCCATCGGCACCG 598
DB 2396 AGGTGAGCAGTAGCAGCATCTATAGAAAATCTGTGACATAAAGCTATTAGGTACAG 2455

QY 599 TGGTGTATGGCCCGCCACCGCTGAACATCATCGCGCAACATGCTGACCCAGCTGGGCT 658
DB 2456 TATTAGTAGACCTACACCTGTACACATTAATTGGAGAATAATCTGTTGACCCAGATCGGCT 2515
QY 659 GCACCTGTAACTTCCCGCATCAGCCCCATCGAGACCGTGCCTGTAAGCTGAAGCCCGGCA 718
DB 2516 GCACCTTGAACCTTCCCGCATCAGCCCTATTGAGACGGTGCCTGTAAGTTGAAGCCGGGA 2575
QY 719 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTACCAGGAGGAAGATCAAGGCCCTTGACCG 778
DB 2576 TGGACGGCCCAAGGTGAAGCAGTGGCCCTTACCAGGAGGAAGATCAAGGCCCTTAGTCG 2635
QY 779 CCATCTCGCAGGAGATGGAGAAGGAGGAGGATCACCAGATCGGCCCGGAGAACCCCT 838
DB 2636 AATCTGTACAGAGATGGAGAAGGAGGAAGATCAGCAAGATCGGCCCTTGAGAACCCCT 2695
QY 839 ACAACACCCCGGTGTCGCATCAAGAAGAAGACAGCACCAGTGGCGCAAGCTGGTGG 898
DB 2696 ACAACACTCCAGTCTTCGCAATCAAGAAGAAGACAGTACCAAGTGGAGAAGCTGGTGG 2755
QY 899 ACTTCCGCGAGCTGAACAAGCCGCCAGGACTTCTGGAGAGTGCAGCTGGGCATCCCC 958
DB 2756 ACTTCAGAGAGCTGAACAAGAACTCAGGACTTCTGGGAAGTTTCTGAGGATCCCCAC 2815
QY 959 ACCCGCGCGCTGAAGAAGAAGAGAGCGTGCACCTGCTGACGCTGGCGGACGCTACT 1018
DB 2816 ATCCCGCTGGTGAAGAAGAGAGTCAAGTGCAGTGTGGTGTGAGTGTGCTACT 2875
QY 1019 TCAGCTGCCCCCTGGACGAGGACTTCCGCAAGTACACCCCTTCCACCATCCCCAGCATCA 1078
DB 2876 TCTCCGCTTCCCTTGGACGAGGACTTCAGGAAGTACACTGCCCTTACGATACCTAGCATCA 2935
QY 1079 ACAACGAGACCCCGCGCATCCGCTACCAAGTACAAAGTGTGCTGCCCGAGGCTGGAAGGCA 1138
DB 2936 ACAACGAGACCCCGCGCATCCGCTACCAAGTACAAAGTGTGCTGCCACAGGATGGAAGGAT 2995
QY 1139 GCGCCAGCATCTTCCAGAGCAGCATGACCAAGATCTGGAGCCCTTCCGCGCCGCAACC 1198
DB 2996 CACGAGCATCTTCAACGAGCATGACCAAGATCCTGGAGCCCTTCCGCAAGCAAAACC 3055
QY 1199 CGGAGATCTGTATCTACCA-----GGCCCCCTGTACGTGGGCGAGCACTTGGAGATCG 1252
DB 3056 CAGACATCTGTATCTATCATAGTACATGGAGGACCTCTACCTAGGAAGTACCTTGGAGATCG 3115
QY 1253 GCAGCAGCCGCGCAAGATCGAGGAGCTCGCAAGCAGCTGCTGCGCTGGGCTTTCACCA 1312
DB 3116 GCGAGCACAGGACCAAGATCGAGGAGCTGAGACAGCTGTGTGAGGTGGGAGTGAACA 3175
QY 1313 CCGCCGACAAAGAACCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1366
DB 3176 CACGAGCAAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3235
QY 1367 CGGCAAGTGGAGCTGCGAGCCCATCGAGCTGCGCGAGAGGAGAGCTGGACCGTGAAGC 1426
DB 3236 CTGACAAGTGGAGCTGCGAGCCCATCGTGTGCTGCTGAGAGGAGCAGCTGGAGTGAAGC 3295
QY 1427 ACATCCAGAGCTGTGGGCAAGCTGAACCTGGCGCAGCAGCATCTACCCCGCATCAAGG 1486
DB 3296 ACATCAGAGCTGTGGGCAAGTGAACCTGGGCAAGCAGCATCTACCCAGCATCAAG 3355
QY 1487 TGGCGCAGCTGTGCAAGCTGTGCGCGGCGCAAGGCCCTGACCCAGCATCTGCGCCCTGA 1546
DB 3356 TTAGCAGCTGTGCAAGCTGTGCGAGGAGCAGGAGCAGTACAGAGTGTATCCACTGA 3415
QY 1547 CGGAGAGCGCAGCTGGAGCTGGCGGAGAACCGGAGATCTCTGCGGAGCCCGTGCAGC 1606
DB 3416 CAGAGGAAGCAGAGCTAGAACTGGCAGAGAACCGGAGATCTCTGAGGAGCAGTACATG 3475
QY 1607 GGTGTACTACCAACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1666
DB 3476 GAGTGTACTACCAACCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3535
QY 1667 AGTGGACCTTACAGATCTTACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1726

Db 3536 AATGGACCTACCAAACTACACAGAGCCCTTCAAGAACCTGAAGACAGGCAAGTACGCCAA 3595
QY 1727 AGATGCCACCGCCACCAACAGACCTGAAGCAGCTGACCGAGGCGGTGCAGAAGATCG 1786
Db 3596 GGATAGGGGTGCCACACCAACAGATGTGAAGCAGCTGACAGAGGCGATGAGAAGATCA 3655
QY 1787 CCATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCCCGCTGCCCATCCAGAGG 1846
Db 3656 CCACAGAGACATCGTGATCTGGGGCAAGACTCCCAAGTTCAAGCTGCCCATACAGAAGG 3715
QY 1847 AGACCTGGGAGACCTGGTGACCGACTACTCTGGCAGGCCACTCGATCCCGAGTGGGAGT 1906
Db 3716 AGACATGGGAGACATGGTGACCGAGTACTCTGGCAAGCCACTGGATCCCTGAGTGGGAGT 3775
QY 1907 TCGTGAACACCCCCCTCGTGAAGCTGTGTACCAAGCTGGAGAGGAGGCCATCATCG 1966
Db 3776 TCGTGAACACCCCTCGTGTGAAGCTGTGTATCAAGCTGGAGAGGAGGCCATCGTGG 3835
QY 1967 GCGCCGAGACCTTCTAGCTGGAGCGCGCCCAACCGCAGACCAAGATCGGCAAGCGG 2026
Db 3836 GAGCAGAGACCTTCTAGCTGGATGGGGCAGCCCAACAGGAGACCAAGCTGGGCAAGSCAG 3895
QY 2027 GCTACGTGACCGACCGGGCCGGCAGAGAAGATCGTGAGCCCTGACCGAGACCAACACAGA 2086
Db 3896 GCTACGTGACCAACCGGAGGACGACAGAAAGTGTGTGACCTGACTGACACCAACCAACAGA 3955
QY 2087 AGACCGAGCTGACGGCCCTGAGCTGGCCCTGAGGACAGCGCAGCGAGTGAACATCG 2146
Db 3956 AGACTGAGCTGCAAGCCATCTACTAGCTCTGCAAGACACCGACTGGAAGTGAACATCG 4015
QY 2147 TGACCGACAGCCAGTACGCCCTGGGCATCATCCAGGCCCGCCGACAGAGCGAGCG 2206
Db 4016 TGACAGACTCACAGTACGCCTGGGCATCATCCAAGCACACAGACCACTCCGAGTCAG 4075
QY 2207 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGGAGAGGTGTACTGAGCTGGG 2266
Db 4076 AGCTGGTGAACAGATCATCGACAGCTGATCAAGAGGAGAAAGTGTACTGGCATGGG 4135
QY 2267 TGCCCGCCCAAGAGGCGATCGGGGCAACAGCAGCATCGACAAAGCTGGTCAGCAAGGGCA 2326
Db 4136 TACCAGCACAAAGGAAATGGAGGAATGAACAAGTAGATAAATAGTCAGTGTCTGGGA 4195
QY 2327 TCCGCAAGTGTCTTCCTGGACGGCATCGAT 2358
Db 4196 TCCGGAAGTGTCTTCTGGACGGATCGAT 4227

RESULT 3

US-09-872-733-1
; Sequence 1, Application US/09872733
; Patent No. US20010036655a1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4338
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Gag/Pol gene
US-09-872-733-1

Query Match 61.2%; Score 1503.8; DB 10; Length 4338;
Best Local Similarity 79.8%; Pred. No. 2.6e-239;
Matches 1896; Conservative 0; Mismatches 447; Indels 34; Gaps 9;

QY 14 TGGCCGAGGCGCATGAGCAGCGCCACCAGC--GCCACATCTCTGATCAGCGCAGCAACT 70
Db 1085 TGGCCGAGGCGCATGAGCAGGTGACGAACCTCGGCGCAGCATAAATGATCAGAGAGCAACT 1144
QY 71 TCRAAGGCCCCCAAGCGCATCATCAAGTGTCTCAACTCGGCAAGAGGGCCACATCGCCC 130
Db 1145 TCCGGAACACCGAGGAAGATGCTCAAGTGTCTCAATTTGGCAAGAGAGGGCACACCGCCA 1204
QY 131 GCAACTCGCCCGCCCCCGCCCAAGAAAGGGCTGCTGGAAGTGGCGCAAGAGAGGGCCACCAGA 190
Db 1205 GGAATGCCCCGGCCCCCGGAGAGAGGGCTGTGGAAATGTGGAAGAGAGGACACCAAA 1264
QY 191 TGAAGGACTGCACGAGCGCCAGGCCAACTTCTTCGCGAGGAGCTGGCCCTTCCCCCAGG 250
Db 1265 TGAAGAGATTGACTGAGAGACAGGCTAA-TTTTATTAGGAAGATCTGGCCCTTCCTCAAG 1323
QY 251 GCRAAGGCCCCGGAGTTCCCGCAGCAGCAGAACCGCCGCAACAGCCGCCACCGCCGCGAGC 310
Db 1324 GGAAGGCCAGGGAAATTTCTTCAGAGCAGACAGCCAGAGCCACAGCCCGCCAGAAAGAGC 1383
QY 311 TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGCGCGCGCGAGCGCGCAGCGGCA 364
Db 1384 TGCAGGTGCTGGGTAGAGACAACAACCTCCGCCCTCAGAAGCAGGAGCGCATAGACAAGAA 1443
QY 365 -----CCCTGAACCTTCCCGCAGATCAACCTGTGGCAGCGCCCTCTGTGAGCATCAAG 418
Db 1444 CTGTATCTTTAACTTCTCCCTCAGATCACTCTTTGGCAACAGCCCTCTGTACAGTAGGA 1503
QY 419 TGGCGCGCAGATCAAGSAGSCCTGCTGGACACCGCGCGCGCAGCAGACCGTGTGGAGG 478
Db 1504 TCGGGGGCAACTCAAGAAAGCGCTGCTCGATACAGAGCAGATGATACAGTATTAGAAG 1563
QY 479 AGATGAGCCCTGCCCGGCAAGTGGAAAGCCCAAGATGATCGGCGCATCGGGGCTTTCATCA 538
Db 1564 AAATGAGTTTGCAGGAAGATGGAACCAAAATGATAGGGGGATCGGGGCTTTCATCA 1623
QY 539 AGTGCCCGCAGTAGCAGCAGATCTGATCGAGATCTCGGCAAGAGGCCATCGGACCG 598
Db 1624 AGGTGAGCGATACGACCATCTATAGAAATCTGTGGACATAAAGCTATAGGTACAG 1683
QY 599 TGTGTATCGCCCCACC-----CCGTCGAACATCATCGCCGCAACATGCTGACCCACGTG 654
Db 1684 TATTAGTAGGACCTTACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACCCAGATC 1743
QY 655 GGCTGCACCCCTGAACCTTCCCGCATCAGCCCATCGAGACCTGCGCGGTGAAGCTGAAGCCC 714
Db 1744 GGCTGCACCTTGAACCTTCCCGCATCAGCCCTATTGAGACGGTGCCTGTTGAAGCCG 1803
QY 715 GGCATGACGGCCCAAGGTGAAGCAGTGGCCCTGTACCGAGGAGAGATCAAGGCCCTG 774
Db 1804 GGGATGACGGCCCAAGGTGAAGCAGTGGCCATTTGACGAAAGAGAGATCAAGGCCCTTA 1863
QY 775 ACCGCCATCTCGAGGAGATGGAAGAGGCGCAAGATCACCAAGATCGCCCCCGAGAC 834
Db 1864 GTCGAAATCTGTACAGAGATGGGAAGGAAGGAAGATCAGCAAGATCGGGCCCTGGAAC 1923
QY 835 CCCTACAACACCCCGCTGTTTCGGCATCAAGAAGAGAGCAGACCAAGTGCAGCAAGCTG 894
Db 1924 CCCTACAACACTCCAGTCTTCGCAATCAAGAAGAGCAGTACCAAGTGGAGAAAGCTG 1983
QY 895 GTGACTTCCGAGCTGAACAAGCGCACCCAGACTTCTGGGAGGTGCAGCTGGGCATC 954
Db 1984 GTGACTTTCAGAGAGCTGAACAAGAGAACTCAGGACTTCTGGGAAGTTCAGCTGGGCATC 2043
QY 955 CCCCACCCCGCCCGCTGCAAGACAAGAGCGCTGACCTGTGGAGCTGGCGACGCC 1014
Db 2044 CCATCCCGCGTGGTGTGAAGAAGAAAGTCACTGACAGTGTGGATGTGGGTGATGCC 2103

```
QY 1015 TACTTACGCTGCCCTTGAGGAGACTTCGCCAAGTACACCCGCTTCACCTCCACG 1074
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2104 TACTTCTCCGTTCCCTTGAGGAGGACTTCAGGAAGTACACTCCCTTCAGTACCTAGC 2163
QY 1075 ATCAACACAGACACCCCGGCATCCGCTACCGATGTAACGTGCTGCCCGCAGGCTGGAG 1134
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2164 ATCAACACAGACACACCGACATCCGCTACCGATGTAACGTGCTGCCCGCAGGAGTGAAG 2223
QY 1135 GGCAGCCCCAGCATCTTCCAGACGACATGACCAAGATCCTGGAGCCCTTCGCGCCCGC 1194
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2224 GGATCACCCAGCATCTTCAAGACAGCATGACCAAGATCCTGGAGCCCTTCGCAAGCAA 2283
QY 1195 AACCCCGAGATCGTATCTACCA-----GGCCCCCCTGTACGTGGCGACGACCTGGAG 1248
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2284 AACCCAGACATCGTATCTATCAGTACATGGACGACCTCTACGTAGGAAGTACCTGGAG 2343
QY 1249 ATC-GGCCACACCGCCCGCATGATCGAGGAGCTGGCGAGACACCTGCTGGCTGGGGCTT 1307
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2344 ATCGGGGCGACAGGACCAAGATCGAGGAGCTTGACACAGCATCTGTTGAGTGGGGACT 2403
QY 1308 CACCACCCCGACAAGACACCAAGAGGAGGCCCTTCCTGCCCAT-----CGAGCT 1361
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2404 GACCACACAGACAGAGACACCAAGAGGAACTCTCTCTGTGATGGCTACGAATC 2463
QY 1362 GCACCCCGCAAGTGGACCGCTGCAGCCCATCGAGCTGCCCGAGAAGGAGAGCTGGACCGT 1421
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2464 GCATCTGACAAGTGGACAGTGCAGCCCATCGTGTGCTGAGAAAGACAGCTGGACTGT 2523
QY 1422 GAAGGACATCCAGAGCTGTGGGCAAGCTGAAGTGAAGTGGCCAGCCAGCATCTACCCGGCAT 1481
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2524 GAAGGACATCAGAAAGTCTGTGGGCAAGTTGAATCTGGCAAGCAGCATCTACCCAGCAT 2583
QY 1482 CAAGTTCGCGCAGCTGTGCAAGTGTGTGCCGGCGCACAGGCCCTCACCGACATCTGCC 1541
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2584 CARAGTTAGCAGCTGTGCAAGTGTCTCAGGAACCAAGGCACTGACAGAAGTATCCC 2643
QY 1542 CCTGACCGAGGAGCGGAGCTGGAGTGTGCCGAGAACCCGAGATCCTGCGGAGGCCGT 1601
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2644 ACTGACAGAGGAAGCAGAGTAGAATCTGGCAGAGAACCCGAGAGATCCTGAAGGAGCACT 2703
QY 1602 GCAGGGGCTACTAGACCCCGACAGGAGACCTGGTGGCCGAGATCCAGAACGAGGCCA 1661
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2704 ACATGGAGTGTACTAGACCCAGCAAGGAGACCTGTATCGCAGAGATCCAGAAGCAGGGCA 2763
QY 1662 CGACAGTGGACCTTACCAGATCTACAGAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTA 1721
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2764 AGCCATGACACTACCAATCTACAGGAGCCCTTCAAGAACCTGAAGACAGGCAAGTA 2823
QY 1722 CGCCAAAGTCCGACCGCCCGACACCAACGAGCTGAAGCAGCTGACCGAGGCGGTGCAGAA 1781
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2824 CGCAAGGATCAGGGTGTGCCACACCAACAGATGTGAAGCAGCTGACAGAGGCAATGCAGAA 2883
QY 1782 GATCGCATGGAGAGATCTGTATCTGGGCAAGACCCCAAGTTCGGCTGCCCATCCA 1841
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2884 GATCACACAGAGAGATCTGTATCTGGGCAAGATCTCCAAGTTCCAAGTTCGCCATACA 2943
QY 1842 GAAGGAGACCTGGAGACCTTGTGGACCGACTACTGCGAGGCCACCTGGATCCCGAGTG 1901
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 2944 GAGGAGACATGGAGACATGTGTGGACCGAGTACTGGCAGGCCCTGGATCCCTGAGTG 3003
QY 1902 GGAGTGTGTAACACACCCCGCCCTGGTGAAGCTGTGGTACCAAGCTGGAGAAGAGGCCAT 1961
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3004 GGAGTTCGTGAACACACCCCTCCCTTGTGTAACCTGTGTATCAGCTGGAAGAAGCCAT 3063
QY 1962 CATCGGGCCGAGACCTTCTACGTGAGCGGCGCCGCAACCGGGAGACCAAGATCGGCAA 2021
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3064 CGTGGGAGCAGACCTTCTACGTGATGGGCGAGCAACAGGGAGACCAAGTGGGCAA 3123
QY 2022 GGCCTGCTAGTGNACCGAGCGGGCGCGGAGAGATCTGTGAGCTCACCGAGACCAACCAA 2081
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3124 GCGAGGCTAGTGACCAACCGAGGAGACAGAAAGTGGTGACCCCTGACTGACACCAACCAA 3183
QY 2082 CCAGAAGACCGAGCTGCAGGGCCATCCAGCTGGCCCTGCAGGACAGCGGAGGAGTGAA 2141
```

```
DB 3184 CCAGAAGACTGAGCTGCRAGCCATCTACTAGTCTGCAAGACAGCGGACTGGAGTGAA 3243
QY 2142 CATCTGACGACACGACAGTACGCTGGCATCATTCAGGCCAGCCAGCCGACAGAGGGA 2201
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3244 CATCTGACAGACTCACAGTAGC-CATGGCATCATCCAAGCACCAACAGACCAATCGA 3302
QY 2202 GAGCAGCTGGTGAACAGATCATCGACGATCATCAAGAGGAGAGGTTACTCTGAG 2261
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3303 GTACAGCTGGTGAACAGATCATCGACGATCATCAAGAGGAGAAAGTGTACTCTGGC 3362
QY 2262 CTGGTGCCGCCCAAGAGGATCGGCGGCAACGAGAGATCGACAAAGCTGGTGAGCAA 2321
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3363 ATGGGTACCAGCACACAAAGAAATGGAGGAAATGAACAAGTAGATAAATAGTCAGTGC 3422
QY 2322 GGGCATCGCAAGGTGCTGTTCTCTGGACGGCATCGAT 2358
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 3423 TGGGATCCGGAAGGTGCTGTTCTCTGGACGGATCGAT 3459
RESULT 4
US-09-872-733-3
; Sequence 3, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-42870S1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872.733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutated Human
; OTHER INFORMATION: Immunodeficiency Virus - 1 Pol gene
US-09-872-733-3
Query Match 46.3%; Score 1137; DB 10; Length 2467;
Best Local Similarity 83.4%; Pred. No. 5e-179;
Matches 1319; Conservative 0; Mismatches 250; Indels 12; Gaps 2;
```

```
QY 790 GAGATGGAGAGGAGGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACAACACCCC 849
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 7 GAGATGGAGAGGAGGGAAGATCAGCAAGATCGGGCTGAGAACCCCTACAACACTCCA 66
QY 850 GTGTTCCGCATCAACAAGAAGGACACCAAGTGGCGCAAGCTGGTGGACTTCGCGAG 909
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 67 GTCTTCGCAATCAAGAAGAGGACAGTACCAGTGGAGGAAAAGCTGGTGGACTTCAGAGAG 126
QY 910 CTGAACAAGCGCACCCAGGACTTCTGGAGGTTGAGTGGGCATCCCCACCCCGCCGCGC 969
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 127 CTGAACAAGAACTCAGGACTTCTGGGAAGTTTCAAGTGGGCATCCCCACATCCCGTGG 186
QY 970 CTGAAGAAGAAGAGCGTGACCGTGTGACGTGGCGAGCGCTTACTTCAGGCTGCC 1029
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 187 TTGAAGAAGAAGAAGTCAAGTGCATGTGTGATGTGGTGTGCTTCTCGCTTCCC 246
QY 1030 CTGGACGAGGACTTCCCAAGTACACCCCTTACCATCCCGACATCAACAAGAGACC 1089
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 247 TTGGACGAGGACTTCAGGAAGTACACTGCCCTTACGATACCTAGCATCAACAAGAGACA 306
QY 1090 CCGGCACTCCGCTACCAAGTACAACTGCTGCCCGAGGCTGGAAGGGCAGGCCCATC 1149
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
DB 307 CCAGGCACTCCGCTACCAAGTACAACTGCTGCCACAGGGATGGAAGGGATCACCAGCCATC 366
```

QY 1150 TTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACCCGAGATCGTG 1209
DB 367 TTTCAAAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCAAGCAAAACCCAGATCGTG 426
QY 1210 ATCTACCA-----GGCCCCCTGTACGTGGGAGGAGCTGTGAGATCGGCCACACCGC 1263
DB 427 ATCTATCAGTACATGAGCAGCACTCTACGTAGGAAGTGACCTGTGAGATCGGGCAGCAGG 486
QY 1264 GCCAAGATCGAGGAGCTGCCAAGCAGCCTCTGCGCTGGGGCTTCACCAACCCCGACAAG 1323
DB 487 ACCAAGATCGAGGAGCTGAGACAGCATCTGTTGAGTGGGGAGCTGACCAACACGAGACAAG 546
QY 1324 AAGCAGCAAGAGGAGCCCTTCCTGCCCAT-----CGAGCTGCCACCCCGACAAGTGG 1377
DB 547 AAGCAGCAAGAGCAACCTCCCTTCCTGTGATGGGTACGAACCTCTGTGACAGTGG 606
QY 1378 ACCGTGAGCCCATCGAGCTGCCGAGAGGAGAGTGTGACCGCTGAACGACATCCAGAAAG 1437
DB 607 ACAGTGCAGCCCATCGTGTCTGCTGAGAGGAGCAGCTGTGTAACGACATACAGAAAG 666
QY 1438 CTGCTGGCAAGCTGAACCTGGCGCAGCCAGATCTACCCGGCATCAAGTGCCTCAGCTG 1497
DB 667 CTGCTGGCAAGTGTGAACCTGGCGCAGCCAGATCTACCCAGGCATCAAAAGTTAGGCACTG 726
QY 1498 TGAAGCTGTCTGCGCGCGCAGGCGCTGACCGACATCTGTGCCCTGACCGAGGAGGCC 1557
DB 727 TGAAGCTGTCTGAGGAAACCAAGGCATGACAGAAAGTGTATCCACTGACAGAGGAAGCA 786
QY 1558 GAGCTGGAGCTGCCGAGAACCCGAGATCTGCGGAGGCGCTGTGACGCGGTGTACTAC 1617
DB 787 GAGCTGAAGCTGCGAGAACCCGAGAGTCTCTGAAGGCGCAGTATGAGTGTACTAC 846
QY 1618 GAGCCAGCAAGCACTGTGTGGCGGAGATCCAGAACGAGGCGCACCAAGTGGACCTAC 1677
DB 847 GAGCCAGCAAGCACTGTGTGGCGGAGATCCAGAACGAGGCGCAAGTGGACCTAC 906
QY 1678 CAGATCTACAGGAGCCCTTCAAGAACCTGAAGACCGGCAAGTACGCCAAGATGCGCACCC 1737
DB 907 CAAATCTACAGGAGCCCTTCAAGAACCTGAAGACCGCAAGTACGCCAAGATGAGGGGT 966
QY 1738 GCCCAGCAACAGCTGTGAAGCAGCTGACCGAGCGCTGAGAAAGATCGCCATGGAGAGC 1797
DB 967 GCCCAGCAACAGCTGTGAAGCAGCTGACAGAGCGAGTGCAGAAAGATCACCAAGAGAGC 1026
QY 1798 ATCGTGTATCTGGGCAAGACCCCAAGTTCGCGCTCGCCATCAGAAAGAGACCTGGAG 1857
DB 1027 ATCGTGTATCTGGGCAAGACTCCCAAGTTCAGCTGCCATACAGAAAGAGACATGGGAG 1086
QY 1858 ACCTGTGACCGCACTACTGGCAGGCGACCTGGATCCCGAGTGGGAGTTCGTGAACACC 1917
DB 1087 ACATGTGAGCCAGTACTGGCAAGCCACCTGGATCCCTGAGTGGGAGTTCGTGAACACC 1146
QY 1918 CCCCCCTGTGAAGCTGTGTACAGCTGTGAGAAAGAGCCCATCATATCGCGCGCCGAGACC 1977
DB 1147 CTTCCCTGTGTGAAGCTGTGTATCAGCTGTGAGAAAGAACCCCATCTGTGGGAGCAGAGACC 1206
QY 1978 TTCTAGCTGTGACGCGCGCAACCCGAGACCAAGATCGGCAAGGCGCGGTACGTGAAC 2037
DB 1207 TTCTAGCTGTGAGCGAGCCACAGGAGAGACCAAGCTGTGGCAAGGAGCTACGTGAAC 1266
QY 2038 GAGCGGGCGCGGCAAGATCTGTGAGCCTGTGACCGAGACCAACCAAGAGACCGAGCTG 2097
DB 1267 AAGCGAGGAGCAGAAAGTGTGTACCTGTGACTGACACCAACCAAGAGACTGTGAGCTG 1326
QY 2098 CAGGCCATCAGCTGTGAGGAGCAGGCGAGGAGTGAACATGTGTGACCGAGCAGC 2157
DB 1327 CAAGCCATCTACCTAGCTGTGAAAGAGCGGAGCTGGAAGTGAACATCGTGTGACAGACTCA 1386
QY 2158 CAGTACGCTGTGGCATCTCCAGGCGCCAGCCGACAGAGCAGGCGAGCTGTGTGAC 2217
DB 1387 CAGTACGCTGTGGCATCTCCAGGACCAACCAAGCAATCCGAGTGTGAGCTGTGTGAC 1446

QY 2218 CAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGGTGCCGCCAC 2277
DB 1447 CAGATCATCGAGCAGCTGATCAAGAAGGAGAGGTGTACCTGAGCTGGGTACCAGCACAC 1506
QY 2278 AAGGGATCGCGGCGACAGCAGATCGACAAAGCTGTGTGAGCAGGCGATCCGCAAGGTG 2337
DB 1507 AAAGGAATTTGGAGAAATGAACAGTAGATAAATTAGTCACTGTGGGATCCGGAAGGTG 1566
QY 2338 CTGTTCTCTGGACGCGCATCGAT 2358
DB 1567 CTGTTCTCTGGACGCGCATCGAT 1587
RESULT 5
US-09-735-487-7
; Sequence 7, Application us/09735487
; Patent No. US20020042679A1
; GENERAL INFORMATION:
; APPLICANT: de BETHUNE, Marie-Pierre
; APPLICANT: HERTOGS, Kurt
; APPLICANT: PAUWELS, Rudi
; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE
; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY
; FILE REFERENCE: 1377-125P
; CURRENT APPLICATION NUMBER: US/09/735,487
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: 09/117,217
; PRIOR FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2601
; TYPE: DNA
; ORGANISM: HIV-HXB2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(492)
; OTHER INFORMATION: gag Polyprotein
US-09-735-487-7
Query Match 44.0%; Score 1081.6; DB 10; Length 2601;
Best Local Similarity 67.9%; Pred. No. 6.4e-170;
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;
QY 14 TGCGCGAGGCCATGAGCCAGGCCACCA---GCGCCACATCTCTGATGTCAGCGCAGCAACT 70
DB 77 TGCTGAAGCAATGAGCCAAAGTAACAAATTCAGCTACCATAATGATGTCAGAGAGGCAATT 136
QY 71 TCAAGGGCCCCAAGCGCATCATCAAGTGTCTCAACTGCGGCAAGGAGGCGCCACATCGCCC 130
DB 137 TTAGGAACCAAGAAGATTTTAAAGTGTTCATTTGTGCAAGAAGGCGCACACAGCCA 196
QY 131 GCAACTGCGCGCGCCCCCGCAAGAAGGCTGTGTGGAAGTGCAGGCAAGAGGGCCACAGA 190
DB 197 GAAATTCAGGGGCCCCCTAGGAAAAAGGGCTGTGGAATTTGGAATTTGGAAGGAAAGGACACAAA 256
QY 191 TGAGGACTTCACCGAGCGCCAGGCCAACTTCCTCCGAGGAGACCTTCCGCTTCCCCCAGG 250
DB 257 TGAAGATTGTACTGAGAGACAGGCTAA-TTTTTTATGGGAAGATCTGGCCCTTCTCTACAG 315
QY 251 GCAAGGCCCGCAGGTTTCCCGAGCGAGCAACCGCGCAACAGCCGCCACCCAGCGCGAGC 310
DB 316 GGAAGGCCAGGATTTTCTTCAGAGCAGACAGACAGCCACAGCCCCACCCAGAGAGAGC 375
QY 311 TGAGGTGCGCGG-----CGAACACCCCGCAGAGCGCGCGCGAGCGCGCAGGCA 364
DB 376 TTCAGGTGCGGGTAGAGACAACAACCTCCCCCTCAGAAGCAGGAGCGCGATAGACAAGAA 435
QY 365 -----CCCTGAACCTTCCCGCCAGATCACCTGTGTGAGCGCCCGCTTCCCGATCAGG 418
DB 436 CTGTATCTTTAACTTCCCTCAGGTCACTCTTTTGGCAACGACCCCTCGTCACAATAAGA 495

; TYPE: DNA									
; ORGANISM: Human immunodeficiency virus type I									
; -09-999-183-1									
Query Match 44.0%; Score 1081.6; DB 10; Length 4307;									
Best Local Similarity 67.9%; Pred. No. 6.6e-170;									
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps 6;									
QY	14	TGGCCGAGGCGCATGAGCAGCGCCACCA---GCGCCAACTCCTGATCAGCGGCAACT	70						
DB	1088	TGGCTGAAGCAATGAGCCAGTACAAATTCAGTACCTACCAATGATCGAGAGGCAATT	1147						
QY	71	TCAAGGCGCCCAAGCGGCATCATCAAGTGTCTCAACTCGCGCAAGGAGGCGCACATCGCC	130						
DB	1148	TTAGGAACCAAGAAAGATTTTAAAGTGTTCATATTTGCGAAAGAGGCGCACACGCCA	1207						
QY	131	GCAACTCGCGCGCCCGCCCAAGNAGGCTGCTGGAAGTGGCGCAAGGAGGCGCACCA	190						
DB	1208	GAATTCGAGGGCCCTTAGGAAAGAGGCTGTTGGAATGTGAAAGGAAGGACACCCAA	1267						
QY	191	TGAAGGACTCACGCGCGCAGGCCAACTTCTTCCGCGAGGACCTGGCCCTTCCCCAGG	250						
DB	1268	TGAAAGATTGCTAGAGACAGGCTAA-TTTTTAGGGAAGATCTGGCCCTTCTACAG	1326						
QY	251	GCAAGGCGCGGAGTTCCCGCAGCGAGCAGAAACCGCGCCAAAGCCCGCAGCGCGAGC	310						
DB	1327	GGAGGCGCAGGGAATTTCTTTCAGAGCAGACAGCCAGAGCCACAGCCCGCCAGAGAGC	1386						
QY	311	TGCAGGTGCGCGG-----CGACAACCCCGCAGCGAGGCGCGCGCGCGCCAGGGCA	364						
DB	1387	TTCAGGTCTGGGTGAGAGCAACAACCTCCCGCCTCAGAAAGCAGGAGCGGATAGACAAGAA	1446						
QY	365	-----CCCTGAACTTCCCGCAGATCACCCCTGTGGCAGCGCCCGCCTGGTGAGCATCAAG	418						
DB	1447	CTGTATCTTTAACTTCCCTCAGTCACTCTTTGCGAAGCCCGCCTCGTCACATAAGA	1506						
QY	419	TGGCGGCCAGATCAAGGAGCGCCTGTGTGACACCGCGCGCGCGCAGCACCGTGTGAGG	478						
DB	1507	TAGGGGGCAACTAAGGAAGCTCTATTAGATACAGGAGCAGATGATACAGTATTAGAAG	1566						
QY	479	AGATGAGCCTGCCCGCAAGTGGAGCCCAAGATGATCGCGCGCATCGCGCGGCTTCATCA	538						
DB	1567	AAATGAGTTTGCCAGGAGATGGAACCCAAATGATAGGGGAAATGGAGGTTTATCA	1626						
QY	539	AGGTGGCCAGTAGCACAGATCCTGTATCGAGATCTGCGGCAAGAGGCGCATCGGCACCG	598						
DB	1627	AAGTAAGACAGTATGATCAGATCTCATAGAAATCTGTGCACATAAAGCTATAGGTACAG	1686						
QY	599	TGCTGATCGCGCCCGCCCGGTGAACATCATCGCGCGCAACATGCTGACCCAGCTGGGCT	658						
DB	1687	TATTAGTAGGACCTACACCTGTCAACATAATTGGAGAAATCTGTTGACTCAGATTGGTT	1746						
QY	659	GCACCTGAACTTCCCATCAGCCCATCGAGACCGTGGCCGTTGAAGCTGAAGCCCGGCA	718						
DB	1747	GCACTTAAATTTTCCCATTTAGCCCTATTGAGACTGTACAGTAAATTAAGCCAGGAA	1806						
QY	719	TGACGCGCCCAAGGTGAAGCAGTGGCCCTGACCGAGGAGAAGATCAAGCGCCTGACCG	778						
DB	1807	TGGATGSCCCAAAAGTTAAACATGGCCATTGACAGAGAAATAAAGACATTAGTAG	1866						
QY	779	CCATCTCGGAGGATGGAGAGGAGGCAAGATCAACAAAGATCGGCGCCCGAGAACCCCT	838						
DB	1867	AAATTTGTACAGAGATGGAAAGAGGAGGAAAATTTCAAAAATTTGGGCGCTGAAAATCCAT	1926						
QY	839	ACAACACCCCGGTTCGCCATCAAGAAGAGCAGCACAAGTGGCCGAAGCTGGTGG	898						
DB	1927	ACAACTCCAGTATTGGCCATTAAGAAAAGACAGTACTAAATGGAGAAAATTAGTAG	1986						
QY	899	ACTTCGCGAGCTGAACAAGCAGCCAGGACTTCTGGGAGGTGCAGCTGGGCTGCCCG	958						
DB	1987	ATTTGAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAATACCAC	2046						
QY	959	ACCCGCGCGCCTGAAGAGAAGAAGCGGTGACCGTGTGGACGTGGGCGACGCCTACT	1018						
DB	2047	ATCCCGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTACTGGATGTGGTGCATATT	2106						
QY	1019	TCAGCGTGCCTTGGACGAGGACTTCCGCAAGTACACCGCTTCACCATCCCCAGCAACA	1078						
DB	2107	TTTCAGTTCCTTTAGATGAAGACTTCAGGAAGTATATGCACTTTACCATTCTAGTATAA	2166						
QY	1079	ACACGAGAGACCCCGGATCCGCTACCAGTACAACGTGCTGCCCGCAGGGCTGGGAAGGCA	1138						
DB	2167	ACAATGAGACACACAGGATTAGATATCAGTACATGTGCTTCCACAGGGATGGAAGGAT	2226						
QY	1139	GCOCACAGCATTTCCAGAGCAGCATGACCAAGATCTCTGGAGCCCTTCGCGCCCGCAACC	1198						
DB	2227	CACGAGCAATATTCCAAAGTAGCATGACAAAAATCTTAGAGCCCTTTAGAAAAACAAATC	2286						
QY	1199	CCGAGATCGTGATCTACCA-----GGCCCCCTGTAGTGGGCGAGCGACCTGAGATCG	1252						
DB	2287	CAGACATAGTATTCTATCAATACATGGATGATTTGTATGTAGGATCTGACTTAGAAATAG	2346						
QY	1253	GCCAGCACCCCGCCCAAGATCGAGGAGCTGGCGCAAGCACCTGCTGCGTGGGGCTTCACCA	1312						
DB	2347	GGCAGCATAGACAAAAATAGAGGAGCTGAGACAACATCTCTTGAGTGGGACTTACCA	2406						
QY	1313	CCCCCGACAAAGACACCAAGAGGAGCCCCCTTCTTGCOCAT-----CGAGCTGCACC	1366						
DB	2407	CACGAGCAAAAAACATCAGAAAGAACCTCCATTCTTTGGATGGTATTGAACCTCCATC	2466						
QY	1367	CCGACAACTGGACCGTCGACGCCCATCGAGCTGCCCCGAGAGGAGAGCTGGACCCGTGAACG	1426						
DB	2467	CTGATAAATGGACAGTACAGCCCTATAGTGTGCCAGAAAAAGACAGCTGGACTGTCAATG	2526						
QY	1427	ACATCCAGAAAGCTGGTGGGCAAGCTGAACCTGGGCGACAGCATCTACCCGGCATCAAG	1486						
DB	2527	ACATACAGAACTTAGTGGGNAATTTGAATTTGGGCAAGTCAGATTTACCCAGGATTAAAG	2586						
QY	1487	TGGCCCGAGCTGTGCAAGCTGCTGCGCGGCGCAAGCCCTGACCGACATCGTGCCCTGA	1546						
DB	2587	TAAGGCAATTTATGTAAACTCTTTAGAGGAACCAAGCACCTAACAGAAAGTAATACCACATA	2646						
QY	1547	CCGAGGAGCGGAGCTGGAGCTGGCCGAGAACCCGAGATCCTCGCGAGCGCCCTGGCAGC	1606						
DB	2647	CAGAAGAGCAGAGCTAGAACTGGCAGAAAAACAGAGAGATTTCTAAAAGACCATAGATG	2706						
QY	1607	CGCTGTACTACGACCCCGACCAAGGACCTGGTGGCGCGAGATCCAGAAAGCAGCGCCACGAC	1666						
DB	2707	GAGTGTATTATGACCCATCAAAGACTTAATAGCAGAAATACAGACGAGCGGGCAAGGCC	2766						
QY	1667	AGTGGACCTACAGATCTACAGAGCGCCTTCAAGAACCTGAAGACCGGCAAGTACGCCA	1726						
DB	2767	AATGGACATATCAAAATTTATCAAGAGCCATTTAAAAATCTGAAAAACAGAAAAATATGCA	2826						
QY	1727	AGATGCGCACCCCGCCACCAACGACGCTGAAGCAGCTGACCGAGGCGCTGCAGAAATCG	1786						
DB	2827	GAATGAGGGTGGCCACACTAATGATGTAACAATTTAACAGAGGCGAGTGCAAAAATAA	2886						
QY	1787	CCATGAGAGCATCTGATCTGGGGCAAGACCCCAAGTTCCGCTCCCATCCACAGAG	1846						
DB	2887	CCACAGAAACCATAGTAATATGGGAAGACTCCTAAATTTAACTCCCATACAAAGG	2946						
QY	1847	AGACTGGGAGACCTGGTGGACCGACTTACTTGGCAGGCGCCCTGGATCCCGGAGTGGAGT	1906						
DB	2947	AAACATGGGAAGACATGGTGGCAGAGTATTGGCAAGCCACTGGATTTCTTCTGAGTGGGAGT	3006						
QY	1907	TCTGTGAACACCCCGCCCTCGTGAAGCTGTGGTACCAGCTGGAGAGAGGAGCCCATCATCG	1966						
DB	3007	TTGTTAATACCCCTCCCTTTAGTGAATTTATGGTACCAGTTAGAGAAAGAACCCCATAGTAG	3066						
QY	1967	GCGCCGAGACCTTCTACGTGAGCGCGCCCGCCCAACCCGAGAGCAAGATCGGAAGCGCG	2026						
DB	3067	GAGCAGAACTTCTATGTAGTAGTGGGCAGCTAACAGGGAGACTAAATTAGGAAGAACAG	3126						
QY	2027	GCTACGTGACCGGACCGGCGCGCAGAAAGATCGTGAGCCCTGACCGGAGACCAACACAGA	2086						

Db	3127	GATATGTTACTAATAGACGAGAACAAAAAGTTGTCCACCTTA	CTGCACACACAACAATCAGA	318
Qy	2087	AGACCGAGCTGCAGGCCATCCAGCTGGCCCTGCAGGACAGCGG	CAGCGAGGTGAACATCG	2146
Db	3187	AGACTGAGTTTACAAGCAATTTATCTAGCTTTTCAGGATTCG	GATTCGGGATTTAGAAGTAAACATAG	3246
Qy	2147	TGACCGACAGCCAGTAGCCCTTGGGCTATCCAGCGCCCGAC	CGCCGACAAAGACGAGCG	2206
Db	3247	TAAACAGACTCACAATATGCATTAGGAATCATTTCAAGCA	CAACCCAGATCAAAGTGAATCAG	3306
Qy	2207	AGCTGTTGAACAGATCATCGAGCAGCTCATCAAGAAGGAGA	GGATGTPACCTGAGCTGGG	2266
Db	3307	AGTTAGTCAATCAAAATAATAGACGATTAAATAAAGAAAG	TCTATCTGGCATGGG	3366
Qy	2267	TGCCCCGCCCAAGAGGCATCGCGGCAACGAGCAGATCGACA	AGCTGTTGAGCAAGGCA	2326
Db	3367	TACCAGCACAAAGGAATTTGGAGGAATGAACAAAGTAGATA	ATAATTTAGTCAGTCTGCGAA	3426
Qy	2327	TCGCGAAGTGCTGTTCTTCGAGCGGCATCGAT	2358	
Db	3427	TCAGGAAAGTACTATTTTATAGATGGAATAGAT	3458	
RESULT 10				
US-09-737-190A-1				
; Sequence 1, Application US/09737190A				
; Patent No. US20020102545A1				
; GENERAL INFORMATION:				
; APPLICANT: Shibuya, Tetsuo				
; TITLE OF INVENTION: A Method for Changing a Target Array, a Method for A				
; TITLE OF INVENTION: A Structure, and an Apparatus, a Storage Medium and				
; TITLE OF INVENTION: Transmission Medium Thereof				
; FILE REFERENCE: JP919990270US1 (14043)				
; CURRENT APPLICATION NUMBER: US/09/737,190A				
; CURRENT FILING DATE: 2000-12-14				
; NUMBER OF SEQ ID NOS: 2				
; SOFTWARE: PatentIn Ver. 2.1				
; SEQ ID NO 1				
; LENGTH: 9719				
; TYPE: DNA				
; ORGANISM: Human iImmunodeficiency virus type 1				
US-09-737-190A-1				
Query Match 44.0%; Score 1081.6; DB 10; Length 9719;				
Best Local Similarity 67.9%; Pred. No. 6.9e-170;				
Matches 1610; Conservative 0; Mismatches 734; Indels 28; Gaps				
Qy	14	TGGCCGAGGCCATGACCGAGGCCACCA---GGCCCAACATCCT	TATGACGCGCAGCAACT	70
Db	1877	TGGCTGAAGCAATGAGCCAAGTAACAAATTCAGCTACCATA	TATGATGCAGAGAGCAATT	1936
Qy	71	TCAAGGSCCCCAAGCGCATCATCAAGTGCTTCAACTGGCG	CAAGGAGGCCACATCGCCC	130
Db	1937	TTAGAACCAAGAAAGATGTTTAAGTGTTCATTTGGCAAG	AGGGCACACAGCCA	1996
Qy	131	GCAACTGCCCGCCCGCCGCAAGAGGGCTGCTGGAAGTGC	CGCAAGGAGGCCACCCAGA	190
Db	1997	GAAATTCAGGGCCCTAGSAAAAAGGGCTGTTGGAATGT	GGAAGGAGACACCCAA	2056
Qy	191	TGAAGGACTGCACGAGCGCCAGGCCAACTTCTCCGGAGAG	ACCTTGCCCTTCCCCGAGG	250
Db	2057	TGAAAGATTGTACTGAGACACAGGCTAA-TTTTTTAGG	GAAGATCTGGCCCTTCTTACAAG	2115
Qy	251	GCAAGGCCCGCGAGTTCGCCAGCGAGCAACCGGCCCA	CAGCCACCGCCGCGGAGC	310
Db	2116	GGAAGCCAGGGAAATTTTCTCAGAGCACCAGAGCCAA	CAGCCACCGCCGCGGAGC	2175
Qy	311	TGCAAGTGC	CGCG-----CGACAAACCCCGCAGCGAGCGCGCG	364
Db	2176	TTCAAGTCTGGGTAGAGACACAACTCCCTCTCAGAGC	AGGACCCGATAGACAAGGAA	2235
Qy	365	-----CCCTGAACCTTCCCCCAGATCACCTCTGGC	ACGCCCCCTTGCTGAGCATCAAG	418


```
QY 1487 TGCCTGAGCTGTGCAAGCTGCTGCGCGCCCAAGGCCCTGACCGACATCGTGCCCTGA 1546
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3376 TAAGGCAATTATGTAACTCTTAGAGGAACCAAGACACTAACAGAAAGTAATACCACTAA 3435

QY 1547 CCGAGAGGCGGAGCTGGAGCTGCGCGAGAACCGGAGATCTCTGCGGAGCCCTGTGACG 1606
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3436 CAGAAGAGCAGAGCTAGAAGCTGCGAGAAAACAGAGAGATTCTAAAAGAACCACTATATG 3495

QY 1607 GCCTGTACTAGACCCAGCAAGGACCTGTGTGGCGGAGATCCAGAACGCGCCACGACC 1666
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3496 GAGTGTATTAGCCCATCAAGACCTTATAGCAGAAATACAGAGCGGAGGCAAGGCC 3555

QY 1667 AGTGAGACTACAGATCTACGAGGAGCCCTTCAAGAAACCTGAAGACCGGCAAGTACGCCA 1726
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3556 AATGGACATATCAAAATTTATCAAGAGCCATTAAANAATCTGAAAACAGGAAATATGCAA 3615

QY 1727 AGATGCCAGCGGCCACACACACGACGTGAAGCAGCTGACGAGGCGGTGAGAAATCG 1786
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3616 GAATGAGGGGTGCCACACTAATGATGTAAACAAATTAACAGAGGCACTGCAAAATAA 3675

QY 1787 CCATGGAGAGCATGCTGATCTGGGCAAGACCCCAAGTTCCGCTGCCCATCCAGAGG 1846
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3676 CCACAGAAAGCATAGTAATATGGGAAAGACTTCTTAATTTAAACTGCCATACAAAGG 3735

QY 1847 AGACTGGGAGACTGCTGACCGGACTACTGGCAGGCGCACTTGATCCCGAGTGGGAGT 1906
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3736 AAACATGGGAACATGCTGGACAGAGTATTGGCAAGCCACCTGGATTCTGAGTGGGAGT 3795

QY 1907 TCGTGAACACCCCTGCTGTAAGCTGTGTTACAGCTGGAGAGGAGGCGCCATCATCG 1966
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3796 TTGTTAATACCCCTTCTTAGTGAATATGTTACCAAGTTAGAGAAAGAACCCATAGTAG 3855

QY 1967 GCGCCGAGACTTCTAGCTGGACGCGCGCCCAACCGCAGACCAAGATCGGCAAGCGC 2026
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3856 GAGCAGAACTTCTATGTAGTGGGCGAGCTAACAGGGAGACTAAATTTAGGAAAGACAG 3915

QY 2027 GCTACGTTACCGACCGGGCGCGAGAGATCGTTGAGCCCTGACCGAGACCAACCAACAGA 2086
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3916 GATATGTTACTAATAGAGGAAGACAAAAAGTTGTCAACCCTAACTGACACAAACATCAGA 3975

QY 2087 AGACCGAGCTGCAGGCATCCAGCTGCGCCCTGACGAGACACGCGAGGAGGTGAACATCG 2146
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3976 AGACTGAGTTACAGCAATTTATCTAGCTTTCGAGGATTCGGGATTAAGAAATAAACATAG 4035

QY 2147 TGACCCGACGACAGTACGCTGCGGCATCATCCAGGCGCCAGCCGACAGAGCGAGAGCG 2206
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4036 TAACAGACTCACATATGCAATAGGAATCATTCAGGCACAAACAGATCAAGTGAATCAG 4095

QY 2207 AGCTGTGAACCGAGATCATCGACAGCTGATCAAGAAAGGAGAGGTGTACCTGAGCTGGG 2266
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4096 AGTTAGTCAATCAAAATATAGAGCAGTTAATAAAAAAGGAAAGGTTCTATCTGCGATGGG 4155

QY 2267 TGCCCCCCCACAGGCGATCGGCGGCACACGACAGATCGACAGCTGGTGAAGCAAGGGCA 2326
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4156 TACCAGCACACAGGAAGAAATTTGAGGAAATCAACAAGTAGATAAATTAGTCAGTCTGGAA 4215

QY 2327 TCAGCAAGTGTCTTCTCGACGCGCATCAT 2358
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4216 TCAGGAAGTACTATTTTAGATGGAATAGAT 4247
```

RESULT 11

```
US-09-943-286-3
; Sequence 3, Application us/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
; plasmid.
; OTHER INFORMATION: plasmid.
US-09-943-286-3
```

```
Query Match 41.9%; Score 1029.2; DB 10; Length 8933;
Best Local Similarity 55.2%; Pred. No. 2.8e-161;
Matches 1329; Conservative 277; Mismatches 738; Indels 64; Gaps 7;
```

```
QY 14 TGGCCGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTCTGATGAGCGCAGCAACT 70
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1200 UGUGAGAGCAAUAGGCCAAGUAAACAAUACAGUACCAUAAUGCAGAGAGGCAAU 1259

QY 71 TCRAAGGGCCCCAAGCGCATCATCAAGTCTCAACTGGGCAAGGAGGCCACATCGCCC 130
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1260 UAGGAACCAAGAAAGAUUGUUAAGUUUUAAGUUUUAAGUUUUAAGUUUUAAGUUUUAAG 1319

QY 131 GCAACTCGCCCGCCCGCCCGCAAGAGGGCTGCTGGAAGTGGCGCAAGGAGGCCACACAGA 190
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1320 GAAAUUGCAGGGCCCCUAGGAAAGAGGCUUGUGAAAUUGUGAAAGAGGAGCACCCAA 1379

QY 191 TGAAGGACTGACCGAGCGCCAGGCCAACTTCTTCCGGGAGGAGCTGCGCTTCCCCCAGG 250
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1380 UGAAGAAGUUGACUGAGAGAGAGAGGCUAA-UUUUUUAGGGAAGAUUCUGGCCUUCUACAAG 1438

QY 251 GCAAGGCCCGCGAGT-----TCCCCAGCG 274
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1439 GGAAGGCCAGGAAAUUUUUUUCAGAGCAGACAGCCACAGCCACCAUUUUUUCAG 1498

QY 275 AGCAGAACCCCGCCCAACAGCCCGCCACAGCGCGAGCTGCGAGTGCGGG-----CGACA 328
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1499 AGCAGACAGAGCCCAACAGCCCGCCACAGAGAGAGCUUCAGUUCUGGUGAGAGACAAC 1558

QY 329 ACCCCCGCAGGAGCGCGCGCGCGCGAGCCAGGCA-----CCCTGAACCTTCCCCCAGA 382
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1559 ACUCCCCUUCAGAGCAGGAGGCGGAGGCGGAGUAGACAAGGAACUGUAUCCUUUAACUUC 1618

QY 383 TCACCTGTGTGACGCGCCCTGTGTGAGCATCAAGGTGGCGCGCAGATCAAGAGAGGCC 442
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1619 UCACUUCUUUGGCAACGACACCCUUCGUCACAAUAAAGUAGGGGGCAACUAAAGAGAGCUC 1678

QY 443 TGCTGGACACCGCGCGCGCAGCACACCCTGTGTGGAGGAGATGAGCCTGCGCGGCAAGTGA 502
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1679 UAUUAGAACAGGAGCAGAUAGUACAGAUUUAAGAAAUAGUUGCCAGGAAGAUUGA 1738

QY 503 AGCCCAAGATGATCGCGGCGCATCGGCGGCTTTCATCAAGGTGCGCGCAGTACGACAGATCC 562
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1739 AACCAAAAUUGAAGGGGAAUUGGAGGUUUUUAUCAAGAAAGACAGAUAGUAGAUAC 1798

QY 563 TGATCGAGATCTGCGGCAAGAGCCATCGCACCGTGTGATCGCGCCCGCCACCCCGTGA 622
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1799 UCAUAGAAAUUCUGUGGACAUAAAGCUAAGUACAGUAAUUAUAGAGGACCUACACCUUCA 1858

QY 623 ACATCATCGGCGCGCAACATGCTGACCCAGCTGGGCTGCACCCCTGAACCTTCCCCATCAGCC 682
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1859 ACAUAAUUGGAAGAAAUUCUGUAGUCUACAGAUUGUUUUCACUUAAUUUUUCCCAUAGCC 1918

QY 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCCGGCATGGAGCGGCCCAAGGTGAAGCAGT 742
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1919 CUAUUUGAGACUGUACCAAGUAAAUUAAAGCCAGGAUUGGCGCCCAAAAGAAUAAACAAU 1978

QY 743 GGCCCTTGACCGGAGGAAGATCAAGGCCCTGACCGCCATCTGCGGAGAGATGAGGAAGG 802
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1979 GGCAUUGACAGAGAGAAAAAUAAAGCAUUAUAGAAAAUUUUGUACAGAAAUAGGAAGG 2038

QY 803 AGGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTTACAACACCCCGGTGTTCCCATCA 862
```


Db 1320 GAAUUGCAGGGCCCCAGGAAAAAGGCGUGUUGGAAAUGUGGAAAGGAAAGGACACACAAA 1379
QY 191 TGAAGGACTGCACGACGCGCAGCCCAACTTCTTCGCGAGGACCTGSCCTTCCCCCAGG 250
Db 1380 UGAAGAUUUAUCAGACAGACAGCUGAA-UUUUUUAGGGAAGAUUCUGGCCUUCUACAAG 1438
QY 251 GCAAGGCCCCGAGT-----TCCCCAGG 274
Db 1439 GGAAGGCCAGGAUUUUUUCAGAGCAGACCAGCCACAGCCCCACCAUUUUCUACG 1498
QY 275 AGCAGAACCGGCGCAACAGAGCCCAACAGCGCGAGCTGCAGGTGCGCGG-----CGACA 328
Db 1499 AGCAGACACAGAGCAACAGAGCCCAACAGAGAGAGAGCAGCUCUAGGUCUGGGUAGAGACAACA 1558
QY 329 ACCCCGACGAGGCGCGCGCGAGCGCCAGGSCA-----CCCTGAATTTCCCCCAGA 382
Db 1559 ACUCCCCUCAGAGCAGGAGCGAUAGACAGGAACUGUAUCCUUUAUCCUCCUCCAGA 1618
QY 383 TCACCCCTGTGGCAGCGCCCTCGTGTGAGCATCAAGGTGGCGGCGCAGATCAAGAGGCCCC 442
Db 1619 UCACUCUUGGCAACGACCCCGUCACAAUAAAGAUAGGGGCGCAACUAAAGGAAGCUC 1678
QY 443 TGTGGACACGCGCGCGAGCAGCACCGTGTGGAGGAGATGAGCCTCCCGCGCAAGTGA 502
Db 1679 UAUUAGAUACAGGAGCAGAUAGACAGAUUUUAGAAGAAUAGUUUGCCAGGAAGUAGA 1738
QY 503 AGCCCAAGATGATCGGCGGCATCGGCGCTTCATCAAGGTGCGCGCAGTACGACAGATCC 562
Db 1739 AACCANAUAUAGGGGAAUUGGAGUUUUUAACAAGAUAGACAGAUAGUACAGAUAC 1798
QY 563 TGATCGAGATCTGGGCAAGAGCCATCGGCACCGTGTGATCGGCGCCCAACCCCGGTGA 622
Db 1799 UCAUAGAAUCUGUGGACAUAAAGCAUAGUAGUACAGAUUUUAGUAGAACCUACCCUGUCA 1858
QY 623 ACATCATCGCGCCCAACATCTGACCCAGCTGGGCTGCACCCCTGAACCTTCCCCATAGCC 682
Db 1859 ACAUUAUUGGAAGAAUUCUGUUGCUCAGAUUGUGUUGCAUUUUUCCCAUAGCC 1918
QY 683 CCATCGAGACCGTGCCTGAAAGTGAAGCCCGCGCATGGAGCGCCCAAGGTGAAGCAGT 742
Db 1919 CUUUGAGACUGUACCAAGUAAAUAUAAAGCCAGGAUUGGAGUAGGCCCAAAAGUUAACA 1978
QY 743 GGGCCCTGACCGAGGAAGATCAAGCCCTGACCGCCATCTCGGAGGAGATGGAGAAG 802
Db 1979 GGCCAUUGACAGAGAAAAAUAAGCAUUAUAGUAGAAUUAUUGUACAGAAUUGGAAAG 2038
QY 803 AGGCAAGATCACCAAGTCGCGCCCGAGAACCCCTACAACACCCCGCTTTCGCCATCA 862
Db 2039 AAGGAAAAUUUCAAUUUGGCGCCUGAGAAUCCACAUAUCCUAGAUUUUGCCAUA 2098
QY 863 AGAAGAGGACAGACCAAGTGGCGCAAGCTGGTGGACTTCCGCGAGCTGAACAAGCGCA 922
Db 2099 AGAAAAAGACAGUACUAAUUGGAGAAAAUUUAGUAGUUUUCAGAGAACUUUAUAGAGAA 2158
QY 923 CCAGAGCTTCTGGAGGTGACCTGGGATCCCGACCCCGCGCGCTTGAAGAAGAGA 982
Db 2159 CUCAAGACUUCUGGGAUUUUAUAGGAUUAUACCAUCCCGCGAGGUGUAAAAAGAAAA 2218
QY 983 AGAGCGTACCGTGTGGAGTGGCGGACGCTTACTTCAAGCTGCCCCCTGAGCAGGACT 1042
Db 2219 AUCAGUACAGUUGGAGUUGGUGUAGCAUUUUUUCAGUUUUUAGUUUAGUAGAGACU 2278
QY 1043 TCCGCAAGTACCGCTTCCACATCCCGAGCATCAACAACGAGACCGCCCGCATCCGCT 1102
Db 2279 UCAGGAAGUUAUCGUAUUUACCAUACUAGUUAUAAACAAGAGACACAGGGAUUAUAG 2338
QY 1103 ACCAGTACACGTGTGCTGCCCGAGGTGGAAGGCGACCCCGACATCTTCCAGAGCAGCA 1162
Db 2339 AUCAGUACAUUGUCUUCACAGGGAUUGGAAGGAUACACCAAGCAUUAUCCAAAGUAGCA 2398
QY 1163 TGACCAAGATCTCGAGCCCTTCCGCGCGCGCAACCCCGAGATCTGTAICTACCA----- 1217
Db 2399 UGACAAAAUUCUAGAGCCUUUUUAAAAACAAAAUCCAGACUAGUUUUCUACAUAACA 2458

QY 1218 -GGCCCCCTGTAGCTGGCGAGCGACCTGGAGATCGGCACGACCGCGCCCAAGATCGAGG 1276
Db 2459 UGGUAGUUAUUGUAGAGUAGUCUAGUAGAAUAGGCGAGCAUAGAACAAAAUAGAGG 2518
QY 1277 AGCTGCGCAAGACCTGCTGCGCTGGGCTTACACACCCCGCAGCAAGAACACCAAGG 1336
Db 2519 AGCUGAGACAAUCUCUGUAGUGGAGCUCUACCAACACAGACAAAAACAUCAGAAAG 2578
QY 1337 AGCCCCCTTCTGCCCCAT-----CGAGTGCACCCCGACAAGTGGACCGTGCAGCCCA 1390
Db 2579 AACCUCAUUCUUGGAGUGGUUAUAGAACUCCAUCCUGAUAAUAGGACAGUACAGCCUA 2638
QY 1391 TCGAGCTGCGCGAGAGAGAGCTGACCGCTGACGAGATCCAGAACTGCTGGCGCAGC 1450
Db 2639 UAGUGUCGAGAAAAAGACAGCUGGACUGUCAAAGUACUACAGAAUGUAGUGGGGAAU 2698
QY 1451 TGAACCTGGCGAGCAGATCTACCCCGCATCAAGGTGCGCAGCTGTGCAAGTGTGTCG 1510
Db 2699 UGAUUGGCGAAGUCAGAUUUACCCAGGAAUUAAGCAUUAUGUAAACUCCUUA 2758
QY 1511 GCGCGCCAGAGCCCTGACCGACATCGTGCCTTGACCGAGGAGCGCGTGGAGTGG 1570
Db 2759 GAGGAACCAAGCACAACAGAAUAAUACCAUACAGAAAGACAGAGCUAAGAACUGG 2818
QY 1571 CCGAGAACCCGAGATCTGCGCGAGCCGCTGACGCGTGTACTACGACCCAGCAAGG 1630
Db 2819 CAGNAACACAGAGAUUUAAGAACACAGUACUAGGAGUUAUUGAACCCCAUCAAAG 2878
QY 1631 ACCTGTGGCGAGATCCAGAGCAGCGCCACGACCTGAGTGGACCTTACAGATCTACCAGG 1690
Db 2879 ACUUAUAGCAGAAUACAGAGCGGCGAAGCCAAUUGGACAUAUCAAUUUAUUAAG 2938
QY 1691 AGCCCTTCAAGAACCTGAACACCGCAAGTACGCCAGATGCCACACCGCCCGCACACCA 1750
Db 2939 AGCAUUUUAAAAUCUGAAACAGGAAAUUAGCAUAGAGGGGUGGCCACACUAAUG 2998
QY 1751 ACCTGTAAGCAGCTGACCGAGCGCGTCAGAGATCGCATGGAGAGCATGCTGATCTGGG 1810
Db 2999 AUGUAAACAAUUAACAGAGCGAGUGCAAAAUAAUACCAAGAAAGCAUAGUAAUUGG 3058
QY 1811 GCAAGACCCCAAGTTCGCGCTGCCCCATCCAGAGAGAGACTGGGAGACCTGTGGACCG 1870
Db 3059 GAAAGCUCUCAAUUUAUAAUACCCAUACAAAAGGAAACAUUGGAAACAUUGGUGGAC 3118
QY 1871 ACTACTGGAGCGCACCTGGATCCCCAGTGGGAGTTCTGTAACACACCCCGCCCTGTGA 1930
Db 3119 AGUUAUGGCAAGCCACUGGAUUCUGAGUGGAGUUUUUAUACCCCUUUUAGUGA 3178
QY 1931 AGCTGTGTACAGCTGGAGAGGAGCCCATCATCGCGCGCGAGACCTTCTACGTGGAGC 1990
Db 3179 AUUAUUGGUACAGUUAAGAAAGAACCCAUAGUAGGAGCAGAAACCUUUAUUAUAG 3238
QY 1991 GCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGCTACGTGACCGACCGGGCGCGC 2050
Db 3239 GGGCAGCUAACAGGGAGACUAAUUAAGGAAAGCAGGAUUAUUAUAAACAAAGAAAGAC 3298
QY 2051 AGAGATCTGTGACCTGACCGAGACCAACCAAGAGACCGAGCTGCAGGCCATCCAGC 2110
Db 3299 AAAAGGUUUGCCCCCUAACACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 3358
QY 2111 TGCCCTTACAGGACGCGCAGGAGGTGAACATCTGTGACCGCAGACAGCAGTACGCCCTGG 2170
Db 3359 UACUUUGCAGGUAUAGAGUUAAGAAUAAACAUAUACAGACUACAUUAUUAUAG 3418
QY 2171 GCATCATCCAGGCCCGCGCAAGAGCGAGCGAGCTGGTGAACCGATCATCGAGC 2230
Db 3419 GAAUUAUUAAGCACAACCAAGAAUUAAGUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3478
QY 2231 AGCTGATCAAGAGGAGAGGTGCTACCTGAGCTGGGTGCGCGCCCGCACAAAGGGCATCGCG 2290
Db 3479 AGUUAUAAAAAGGAAAGGCUUUCUGGCAUUGGUGGACGACGACACAAAGAAUUGGAG 3538

```
QY 2291 GCAACGACGACATCCACAAGCTGCTGAGCAGGCGCATCCGCAAGGTGCTGTTCTCTGACG 2350
Db 3539 GAAAGAACAGUAGUAAUAGUCAGUGCGUGGAUCAGSAAAUACUAAUUUAGUAG 3598

QY 2351 GCATCGAT 2358
Db 3599 GAAUAGAU 3606

RESULT 13
US-09-943-286-9
; Sequence 9, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the IAC-Bscr pseudo target
; NAME/KEY: mutation
; LOCATION: (4140)...(4159)
; OTHER INFORMATION: Mutated positions: 4140-42, 4145-47, 4152,
; OTHER INFORMATION: 4156-57, 4159
US-09-943-286-9

Query Match 41.9%; Score 1029.2; DB 10; Length 8933;
Best Local Similarity 55.2%; Pred. No. 2.8e-161;
Matches 1329; Conservative 277; Mismatches 738; Indels 64; Gaps 7;

QY 14 TGGCGGAGGCGCATGAGCCAGG---CCACCAGCGCCACATCTCTGATGCGCGCAGCAACT 70
Db 1200 UGGCUGAAGCAUUGAGCAAGUACAACAAUACAGCUACCAUUAUGAUGCAGAGAGGCAUU 1259

QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTGCGGAGGAGGCGGCACATCGCC 130
Db 1260 UUAGGAACCAAGAAAGAAUGGUUUAUGUUAUGUUAUGGCAAGAAAGAGGCGCACAGCCA 1319

QY 131 GCAACTGCCGCGCCCGCCGCAAGAGGGCTGTGGAAGTGGCGCAAGSAGGCGCACAGA 190
Db 1320 GAAUUGCAGGCGCCCUAGGAAAGAGGCGUGUUGGAAUUGUGGAAAGAGGACACCAAA 1379

QY 191 TGAAGGACTGCACCGAGCGCCAGCGCAACTTCTTCCGCGAGGACCTGGCCCTTCCCCCAGG 250
Db 1380 UGAAGAUAUUGACUGAGAGACAGGCUAA-UUUUUUAGGGAAGAUUCUGCGCCUUCUACAAG 1438

QY 251 GCAAGGCGCGGAGT-----TCCCCAGCG 274
Db 1439 GGAAGCGCAGGGAUUUUUUUUCAGAGCAGACACAGAGCAACAGCCCAACCAUUUUCAG 1498

QY 275 AGCAGAACCGCGCAACAGCGCCACAGCGCGAGCTGCGAGTGCAGGTGCGCGG-----CGACA 328
Db 1499 AGCAGACCAAGCCACAGCGCCCAACAGAGAGAGCUUUCAGGUCUGGUGGAGAGACAACA 1558

QY 329 ACCCGCGCAGCGAGCGCGCGCGCGCGCGCGCA-----CCCTGAATCTCCCCCAGA 382
Db 1559 ACUCCCCCAGAGCAGGAGGCGGAUAGACAAGCAAGCAAGUUAUUCUUUUAUUCUCCUCAGA 1618

QY 383 TCACCCTGTGGCAGCGCCCTCGTGAAGTCAAGTGGCGGCGCGCAGATCAAGAGGCGCC 442
Db 1619 UCACUCUUUGGCAACGACCGCCUGCACAUAUAAAGAUAGGGGCGCAACUAAGAGAGCUC 1678

QY 443 TGCTGGACACCGCGCGAGCACACCTGCTGGAGGAGATGACCTGCGCGCGCAAGTGA 502
Db 1679 UAUUAGUACAGGAGCAGAUAGUUAAGAAUUAAGAAUUAUGUUUGCCCGAGGAAGUAGA 1738
```

```
QY 503 AGCCCAAGATGATCGGGCGCATCGCGCGTCTATCAAGGTGCGCCAGCTACGACAGATCC 562
Db 1739 AACCAAAAUAUGAGGGGAAUUGAGGUUUUAUCAAAAGUAGACAGAUUAUGAUCAGAUAC 1798

QY 563 TGATCGAGATCTCGGGCAAGAGCCATCGGCACCGTCTGATCGGCGCCCAACCCCGTGA 622
Db 1799 UCAUAGAAAUCUGUGGACAUAAAGCAUAGUACAGUAUUAAGAGCAUAGGACCUACACCCUGUCA 1858

QY 623 ACATCATCGCGCCGCAACATGCTGACCCAGCTGGGCTGACCCCTGACCTTCCCATCAGCC 682
Db 1859 ACAUAAUUGGAAGAAUUCUGUUGUCACUAGUUGGUCACUUUAAUUAUUUCCAUUAGCC 1918

QY 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCGCGCATGAGCGCGCCCAAGGTGAAGCAGT 742
Db 1919 CUUUGAGACUGUACAGUAAAUUAAGCCAGAAUUGAUGGCGCCCAAAAGUUAACAAU 1978

QY 743 GGCCCTGACCGAGGAGAGATCAAGCGCTGACCGCATCTGCGAGGAGATGGAGAGG 802
Db 1979 GGCCAUUGACAGAGAAAAAUAAGCAUUAUAGAAAUUUGUACAGAAUUGGAAAGG 2038

QY 803 AGGCAAGATCACCAAGATCGCGCCCGCAGAACCCCTACACACCCCGTGTTCGCCATCA 862
Db 2039 AAGGAAAUUUUCAAUUAUUGGCGCUGAGAAUCCAUACAUAUACUCCAGUAUUUGCCAUAA 2098

QY 863 AGAAGAAGGACAGCACCAAGTGGCGCAAGCTGTGGACTTCCGCGAGCTGAACAGCGCA 922
Db 2099 AGAAAAAGACAGUACUAAUUAUGGAAAUUAUAGUAAUUAUUAUUAUUAUUAUUAUUA 2158

QY 923 CCCAGGACTTCTGGGAGGTGCGAGTGGGCATCCCCACCCCGCGCGCTGAAGAAAGA 982
Db 2159 CUCAAGACUUCUGGGAAGUUAUUAUAGAAUACCAUCCCGCAGGUAUAAAAAGAAA 2218

QY 983 AGACGCTGACCGTCTGACGTGGCGACGCTTACTTCAGCGTGGCCCTGACGAGGACT 1042
Db 2219 AAUCAGAAACAGUACUGGAGUGGUGAUGCAUUAUUUUUUCAGUUUCCUUAUAGAAAGACU 2278

QY 1043 TCCCAAGTACACCGCTTACCATCCCGCATCAACAACGAGACACCCCGCATCCGCT 1102
Db 2279 UCAGAGAUUAUCUGCAUUAUACUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2338

QY 1103 ACCAGTACAGCTGCTGCCCGAGGCTGGAAGGCGAGCCCGCAGCATCTTCCAGAGCAGCA 1162
Db 2339 AUCAGUACAUAUGUUCUCCACAGGGAUGGAAAGGAUACACAGCAUUAUUCAAAGUAGCA 2398

QY 1163 TGACCAAGATCTCGAGCGCTTCCCGCGCCCAACCCCGAGATCGTGATCTACCA----- 1217
Db 2399 UGACAAAUUUUAGAGCCUUUUUAAAAAUAUCCAGACAUAGUUAUUAUUAUUAUUAUUA 2458

QY 1218 -GGCCCCCTGTAGCTGGGCGAGCGACCTGGAGATCGGCGCAGCAGCGCGCCCAAGATCGAGG 1276
Db 2459 UGGAUAGUUUAUUGAUGAGGUAUCUAGUAAUUAUAGGCGCAGUAGCAACAAUUAUAGG 2518

QY 1277 AGTGGCGAAGACCTGCTGGGCTGGGCTTTCACACCCCGCAGCAAGAACACAGAGG 1336
Db 2519 AGCUGAGACAACAUCUGUGAGGUGGAGCUUACACACAGCAACAAAAAUAUUAUUAUUA 2578

QY 1337 AGCCCCCTTCTCCCTCAT-----CGAGCTGACCCCGCAGCAAGTGGACCTTGAGCCCA 1390
Db 2579 AACCUCAUCCUUUGAUGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2638

QY 1391 TCGAGCTGCCCGAGAAGAGAGCTGGACCGTGAACGACATCCAGAGCTGTGGGCAAGC 1450
Db 2639 UAGUGUGCCAGAAAAAGACAGUGGACUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2698

QY 1451 TGAAGTGGCGAGCGCATCTACCCCGCATCAAGGTGCGCGAGCTGTGCAAGTGTGCTG 1510
Db 2699 UGAUUGGGCAAGCAGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2758

QY 1511 GCGCGCCAGAGCGCTGACCGACATCTGTCGCCCTGACCGAGGAGCGCGAGCTGAGCTGG 1570
Db 2759 GAGAAACCAAGCACUAAACAGAGUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2818

QY 1571 CCGAGAACCGCGAGATCTCTGCGGAGCGCGTGTACTACTAGACCCCGAGGAGG 1630
```

[illegible]

RESULT 14

```

US-09-798-675-4
; Sequence 4, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798,675
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186,364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251,083
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4

```

QY 623 ACATCATCGCGCGGAAACATGCTGACCCAGCTGGGTGACCCCTGAACCTTCCCATCAGCC 682
DB 1852 ACATAATTTGAAGAAATCTGTTGACTCAGATGGTTGCACITTAATTTTCCCATAGCC 1911
QY 663 CCATCGACACCGTGGCGGTGAAGCTGAAGCCGGCATGAGCGGCCCCCAAGGTGAAGCACT 742
DB 1912 CTATTTGACAGTGTACCACTTAAATTTAAAGCCAGGAATGATGGCCCAAAAGTTAAACAAT 1971
QY 743 GCGCCCTGACCGGAGGAGATCAAGGCCCTGACCGCCATCTGCGAGGAGATGGAGAAGG 802
DB 1972 GGCATTGCACAGAGAAAATAAAGACATTTAGTAGAATTTGTACAGAAATGGAAAAGG 2031
QY 803 AGGCAAGATCACCAAGATCGGCCCGGAGAACCCCTACAACACCCCGCTGTTTCGCCATCA 862
DB 2032 AAGGAAAAATTTCAAAAATTTGGCGCTGAGAAATCCATACAAATATCCACTATTTGGCAATA 2091
QY 863 AGAAGAGACAGACACCAAGTGGCGCAAGCTGGTGACCTTCGGGAGCTGAACAGCGCA 922
DB 2092 AGAAAAAGACAGTACTAATGGGAGAAATTTAGTAGATTTTCAGAGAACTTAATAGAGAA 2151
QY 923 CCCAGGACTTCTGGAGGTGCGAGCTGGGCATCCCGACCCCGCGCTGGAAGAGAAGA 982
DB 2152 CTCAGACTTCTGGAGTTCATTTAGGAATACCACTCCCGCAGGGTTAAAAAGAAAA 2211
QY 983 AGAGGTGACCGTGTGAGCTGGCGACGCTACTTTCAGCGTGGCCCTGGACGAGGACT 1042
DB 2212 AATCAGTAACAGTACTGGATGGGTGATGCATATTTTCAGTTCCTTTAGTAGAAGACT 2271
QY 1043 TCCCAAGTACACCGCTTCCACCTCCCGACATCAACAGACGACCCCGCGCTGGAAGAGAAGA 982
DB 2272 TCAGAGATGATGCTTCCACAGGATGGAAGGATCACCAGCAATATTCACAGTAGCA 2331
QY 1103 ACCAGTCAACGCTGCTGCCCGAGGCTGGAGGCGACCCCGAGCATCTTCAGAGACAGA 1162
DB 2332 ATCAGTACATGCTTCCACAGGATGGAAGGATCACCAGCAATATTCACAGTAGCA 2391
QY 1163 TGACCAAGATCTCGAGGCTTCCCGCGCCCGCAACCCCGAGATCGTGATCTACCA - - - - 1217
DB 2392 TGACAAAATCTTAGAGCCCTTTTAAAAACAAAATCCAGACATAGTTATCTATCAATACA 2451
QY 1218 -GGCCCCCTCTAGCTGGGACGCGCTGGAGATCGCGACGACCCCGCGCGCAAGATCGAG 1276
DB 2452 TGAACGATTTGTATGTAGGATCTGACTAGNAATAGGCGCATAGAACAAAATAGAGG 2511
QY 1277 AGCTGGCGAAGCACTGCTGGCTGGGGCTTCAACCCCGCGCAAGAACGACCAAGG 1336
DB 2512 AGCTGAGACAACTCTGTTAGTGGGACTTACCACACGACAAAACATCAGAAAG 2571
QY 1337 AGCCCCCTTCTGCCCCAT - - - - -CGAGCTGCACCCCGACAGTGGACCGTGCAGCCCA 1390
DB 2572 AACCTCCATCTTGGATGGGTATGAATCCATCTCCATGATATGAATGGACAGTACAGCCTA 2631
QY 1391 TCGAGCTGCCGGAAGAGAGCTGGACCGTGAACGACATCTCAGAAGCTGGTGGCGAAGC 1450
DB 2632 TAGTGCTGCCAGAAAAGACAGCTGGACTGTAATGACATACAGAAGTTAGTGGGGAAT 2691
QY 1451 TGAATCGGCGACCGCAATCTACCCCGGCATCAAGGTGCGCGAGCTGTGCAAGCTGTGC 1510
DB 2692 TGAATCCGGAAGTCAGATTTACCCAGGGATTTAAAGTAAGGCAATTTATGTAACCTCTTA 2751
QY 1511 GCGSCGCAAGGCCCTGACCGACATCTGCTCCCTGACCGAGGAGCGCGAGTGGAGCTGG 1570
DB 2752 GAGAACCAAGGACCTAACAGAAGTAAATACCCTAACAGAGAGAGCAGAGTAACTGG 2811
QY 1571 CCGAGACCGCGAGATCTCGCGAGCGCTGCGAGCGGTGCTACGACCCCGCGCAAGG 1630
DB 2812 CAGAAACACAGAGATTTCTAAAGAACCACTACATGAGTGTATTTATGACCCATCAAAAG 2871
QY 1631 ACCTGGTGGCGGAGATCCAGAACGAGGCGCACGACCAAGTGGACCTACAGATCTACAGG 1690
DB 2872 ACTTAATAGCAGAAATACAGACGAGGCGCAAGGCCAATGGACATATCAAAATTTATCAAG 2931

QY 1691 AGCCCTTCAAGAACCTGGAAGACCGGCAAGTACGCCNAGATGCGCACCCGCGCACCAACG 1750
DB 2932 AGCCATTTAAATCTGAAAACAGGAAAATATGCAAGATGAGGGGTGCCCACTAATG 2991
QY 1751 AGCTGAAGCAGCTGACCGAGGCGGTGCAAGATCGCCATGGAGAGCATCGTGATCTGG 1810
DB 2992 ATGTAACAACATTAACAGAGGCGAGTCAAAAATAACACAGACATAGTAATATGG 3051
QY 1811 GCAAGACCCCAAGTTCGCCCTGCGCATCCGAAGAGACCTGGGAGACCTGGTGGACG 1870
DB 3052 GAAGAGCTCTTAATTTAAACTACCCATACAAAAGGAAACATGGGAAACATGGTGACAG 3111
QY 1871 ACTACTGCGAGCGCACCTGGATCCCGAGTGGAGTCTGTAACACCCCGCTGGTGA 1930
DB 3112 AGTATTTGCAAGCCACTGGATTCCTGAGTGGAGTTTGTAAATACCCCTCTTATGTA 3171
QY 1931 AGCTGTGTACAGCTGGAGAAGAGCCCATCATCGCGCCGAGACCTTCTACGTGGAG 1990
DB 3172 AATTATGTTACCAGTTAGAGAAAGAACCCATAGTAGGAGCAGAAACCTTCTATGTAG 3231
QY 1991 GCGCGCCCAACCGCGAGACCAAGATCGGCAAGCGCGGTACGTACCGACCGGCGCGG 2050
DB 3232 GGGCAGCTAACAGGGAGACTAAATTTAGGAAAAGCAGGATATGTTACTAACAAAGGAAG 3291
QY 2051 AGAAGTCTGAGCGCTGACCGAGACCAACCAAGCAGACCGAGCTGCGAGCCATCCAG 2110
DB 3292 AAAAGTTGTTCCCTTAATACTAACACAAATCAGAAAACCTCAGTTACAAGCAATTTAT 3351
QY 2111 TGGCCCTCGAGACAGCGCGAGGTGAACATCGTACCGACAGCGAGTACGCCCTGG 2170
DB 3352 TAGCTTTGAGGATTCAGGATTTAGAAAGTAAACATAGTAACAGACTCAACAATGATG 3411
QY 2171 GCATCATCCAGCGCCAGCGCCGACAGAGCGAGAGCGAGTGTGAACAGATCATCGAG 2230
DB 3412 GAATCTTCAAGCACACCAACAGATAAAGTGAATCAGAGTTAGTCAATCAATAATAGAG 3471
QY 2231 AGCTGATCAAGAAAGGAAGGTGTACCTGAGTGGGTGCGCCCGCACAAAGGCGCGG 2290
DB 3472 AGTTAATAAAAGGAAAGGTCTATCTGGCATGGGTACCGACACAAAGGAATTTGGAG 3531
QY 2291 GCACGAGCAGATCGACAAGCTGGTGAAGAGGCGATCCGCAAGGTGTGTTCTCTGGAG 2350
DB 3532 GAAATGAACAAGTAGATAAATAGTCACTGCTGGAATCAGGAAATACTATTTTTAGATG 3591
QY 2351 GCATCGAT 2358
DB 3592 GAATAGAT 3599

RESULT 15
US-09-798-675-5
; Sequence 5, Application US/09798675
; Patent No. US20020106798A1
; GENERAL INFORMATION:
; APPLICANT: Emory University
; TITLE OF INVENTION: HIV VACCINES
; FILE REFERENCE: E056 2020
; CURRENT APPLICATION NUMBER: US/09/798,675
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/186,364
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 60/251,083
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 9918
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Construct of vaccine vector pGal and vaccine insert expressing
; OTHER INFORMATION: ade B HIV-1 Gag-Po
; NAME/KEY: misc_feature

LOCATION: (106)..(1641)
OTHER INFORMATION: encodes proteins of viral inner core
NAME/KEY: misc_feature
LOCATION: (1401)..(3617)
OTHER INFORMATION: encodes viral proteins but not integrase
NAME/KEY: misc_feature
LOCATION: (3708)..(5715)
OTHER INFORMATION: regulates high-level production of HIV genes
NAME/KEY: misc_feature
LOCATION: (3847)..(5944)
OTHER INFORMATION: encodes protein regulating the transfer of RNA to cytoplasm
NAME/KEY: misc_feature
LOCATION: (5939)..(4181)
OTHER INFORMATION: gene participates in viral assembly and budding
NAME/KEY: misc_feature
LOCATION: (4099)..(4941)
OTHER INFORMATION: encodes truncated form of viral coat protein
NAME/KEY: misc_feature
LOCATION: (6098)..(9918)
OTHER INFORMATION: vaccine vector pCAL

US-09-798-675-5

Query Match 41.6%; Score 1022.8; DB 10; Length 9918;
Best Local Similarity 66.5%; Pred. No. 3.2e-160;
Matches 1602; Conservative 0; Mismatches 742; Indels 64; Gaps 7;
QY 14 TGGCCGAGCCATGAGCCAGG---CCACCAGGCGCAACATCTCTGATCGAGCGCAGCAACT 70
DB 1193 TGGCTGAAGCAATGAGCCAGTACAAATACAGCTACCAATGATCGAGAGGCAATT 1252
QY 71 TCAAGGCGCCCAAGCGCATCATCAAGTGTCTCAACTCGGCGAGGAGGCGCACATCGCCC 130
DB 1253 TTAGGAACCAAGAAAGATGGTAAAGAGCTTCAATACGGCGAAGAGGCGCACACGCCA 1312
QY 131 GCAACTCGCGCGCCCGCCGCAAGAGGCTGTCTGGAAGTGGGCAAGAGGCGCCACACAGA 190
DB 1313 GAAATTCAGGGCGCCCTAGGAAAAGGCGAGCTGGAAGAGCGGAAAGAGGACACAAA 1372
QY 191 TGAAGAGTCGACGAGCGCGGCGAGCCAACTTCTTCGCGAGGAGCTTGGCTTCCCGCAGG 250
DB 1373 TGAAGATTGTACGAGAGACAGGCTAA-TTTTATAGGAGATCTGGCTTCTTACAG 1431
QY 251 GCAAGGCGCGGAGT-----TCCCCAGCG 274
DB 1432 GGAAGCGCAGGAATTTTCTTCAGAGCAGACAGAGCCACAGCCCGCCACCATTTCTTCAG 1491
QY 275 AGCAGACCGCGCCCAACAGCCCGCCAGCGCGAGCTGCGAGTGGCGG-----CGACA 328
DB 1492 AGCAGACGAGCCCAACAGCCCGCCAGAGAGAGCTTCAGGTCTGGGGGTAGAGACAACA 1551
QY 329 ACCCGCGAGCGAGCGCGCGCGAGCGCCAGGCA-----CCCTGAACCTTCCCGCAGA 382
DB 1552 ACTCCCCCTCAGAGCAGGAGCGCGATAGACAAGAACTGTATCCTTTAACTTCCCTCAGA 1611
QY 383 TCACCCCTGTGCGACGCGCCCTGTGTAGCATCAAGGTGGGCGCGCAGATCAAGAGGCGCC 442
DB 1612 TCACTCTTTGCAACAGACCCCTCTGCAATAAAGATAGGGGGCACTAAAGGAAGCTC 1671
QY 443 TGCTGGACACCGCGCGCGAGCAGCCGTGTGAGGAGATGAGCTTCCCGCGCAAGTGA 502
DB 1672 TATTAGATACAGGAGCAGATGATACAGTATTAGAGAAATGAGTTTGGCCGGAAGATGGA 1731
QY 503 AGCCCAAGATGATCGGGGCGCTCGCGGCTTCATCAAGGTGCGCGAGTACGACACAGATCC 562
DB 1732 AACCAAAATGATAGGGGAATTGGAGTTTATCAAGTAAGACAGTATGATCAGATAC 1791
QY 563 TGATCGAGATCGCGGCAAGAGGCCATTCGCAACCGTGTGATCGGCGCCACCGCCCGTGA 622
DB 1792 TCATAGAAATCTGTGACATAAAGCTATAGGTACAGTATTAGTAGGAGCTACACCTGTCA 1851
QY 623 ACATCATCGGCGCGCAACATGTCGCCAGCTGGGCTGCACCTGACCTTCCCATCAGCC 682
DB 1852 ACATAATTGGAAGAAATCTGTGTACTCAGATTGGTTGCACTTAAATTTTCCCATTAGCC 1911

QY 683 CCATCGAGACCGTCCCGTGAAGCTGAAGCCCGCATGGACGCGCCCAAGGTGAAGCAGT 742
DB 1912 CTATTGAGACTGTACCAGTAAATTAAGCCAGGAATGGATGCCCCAAAGATTAAACAAT 1971
QY 743 GGCCCTGTACCGAGGAGAGATCAAGGCCCTGACCCGCATCTCGGAGAGATGAGAGAGG 802
DB 1972 GGCCATTGACAGAAGAAAAATAAAGCATTAGTAGAAATTTGTACAGAAATGGAAGG 2031
QY 803 AGGCGAAGATCACCAAGATCGGCCCGGAGAACCCCTACACACCCCGCTGTCGCCATCA 862
DB 2032 AAGGGAATAATTTCAAAATTTGGGCTGAGAAATCAATACTTCCAGTATTTGCCATAA 2091
QY 863 AGAAGAGGAGCAGCACCAAGTGGCGCAAGCTGTGGTGGACTTCCGCGAGCTTCAACAAGCGCA 922
DB 2092 AGAAAAAGACAGTACTTAATGGAGAAAAATAGTAGATTTCAGAGAACTTAATAAGAGAA 2151
QY 923 CCAGAGTCTTCTGGAGGTGCAGCTGGGCTGCCCCACCGCGCGGCTGAAGAGAGAA 982
DB 2152 CTCAGAGCTTCTGGGAAGTTCAATTTAGGAATACCACATCCCGCAGGTTAAAAAAGAAA 2211
QY 983 AGACGCTGACCGTCTGGAGCTGGCGCAGCGCTACTTCAGCGTCCCGCTGGACGAGACT 1042
DB 2212 AATCAGTAACAGTACTGGATGTGGGTGATGCATATTTTCAGTTCCCTTAGATGAAGACT 2271
QY 1043 TCCGCAAGTACACCGCTTCCACCATCCCGCAGCATCAACAGCAGACCGCCCGGCTCCGCT 1102
DB 2272 TCAGGAAGTACTATGCATTTACCATACCTAGTATAACAATAGACACACCGGATTAGAT 2331
QY 1103 ACCAGTACACAGCTGCTGCCCGAGGCTGGAGGCGCAGCCCGCAGCATCTTCAGAGCAGCA 1162
DB 2332 ATCAGTCAATGTCTTCCACAGGGATGGAAGGATCACCAGCAATATTCNAATAGCA 2391
QY 1163 TGACCAAGATCTCGAGCGCTTCCCGCCCGCAACCGCGAGATCGTGATCTACCA----- 1217
DB 2392 TGACAAAAATCTTAGAGCTTTTAAAAACAATAATCCAGACATAGTTATCTATCAATACA 2451
QY 1218 -GGCCCCCTCTAGCTGGGCGAGCCCTGGAGATCGGGCGACCGCGCGCAGAGTCTGAGG 1276
DB 2452 TGAACGATTTGTATGTAGGATCTGACTTAGAAATAGGCGAGCATAGAACAAAAATAGAGG 2511
QY 1277 AGCTGGCGACGACCTGCTGGCTGGGCTTCCACCCCGCGACCAAGAGCAGCAGAGG 1336
DB 2512 AGCTGAGACAACTCTGTTGAGGTGGGAGCTTACCACACAGCAAAAAAATCAGAAAG 2571
QY 1337 AGCCCCCTTCTGCCCCAT-----CGAGCTGCACCCCGCAAGTGGACCTGCGAGCCCA 1390
DB 2572 AACCTCCATCTCTTGGATGGGTTATGAACTCCATCTCTGATAAATGGACAGTACAGCCTA 2631
QY 1391 TCGAGCTGCCGGAAGAGAGAGCTGGCCGTGAGAGCATCCAGAGCTGGTGGGCAAGC 1450
DB 2632 TAGTGTCTGCCAGAAAAAGACAGCTGGAGTCTCAATGACATACAGAAGTTAGTGGGAAAT 2691
QY 1451 TGAAGTGGCGCAGCAGATCTACCCCGCATCAAGTGGCGCAGCTGCAAGCTGTGCTGC 1510
DB 2692 TGAATCGCAGAGTACAGATTTACCCAGGGATTAAGTAGGCAATATTTAACTTCCCTTA 2751
QY 1511 CGCGCGCAAGCGCTGACCGACATCTGTCGCCCTGACCGAGGAGCGCGAGCTGGAGCTGG 1570
DB 2752 GAGGAACCAAGCAGTACAGAAATTAACCACTAACAGAGAGAGCAGAGCTAGAACTGG 2811
QY 1571 CCGAGAACCGGAGATCTCGCGGAGCGCGCTGACGGGTGTACTACACCCCGCAGAGG 1630
DB 2812 CAGAAAAACAGAGAGATTCTAAAGAACCCAGTACATGGAGTGTATTATGACCCATCAAAAG 2871
QY 1631 ACCTGGTGGCGAGATCCAGAGCAGGCGCCAGCAGCTGGACCTTACCAGATCTACCAGG 1690
DB 2872 ACTTAATAGCAGAAATACAGAGCAGGCGGCAAGGCAATGGACATATCAATTTATCAAG 2931
QY 1691 AGCCCTTCAAGAACCTGAAGACCGCGCAAGTAGCCCAAGATGCGCACCGCCCGCAGCAAG 1750
DB 2932 AGCCATTTAAATACTGAAACAGAGAAATATGCAAGAAATGAGGGGTGCCCCACACTAATG 2991

